



TECHNISCHE
UNIVERSITÄT
DARMSTADT

Semantisches Tagging zur Verwaltung von webbasierten Lernressourcen
Modelle, Methoden und eine Plattform zur Unterstützung Ressourcen-basierten
Lernens

Vom Fachbereich
Elektrotechnik und Informationstechnik
der Technischen Universität Darmstadt
zur Erlangung des Grades eines
Doktor-Ingenieurs (Dr.-Ing.)
genehmigte

Dissertationsschrift

von

Doreen Böhnstedt, M. Sc.

Geboren am 06. Oktober 1981 in Eilenburg

Vorsitz: Prof. Dr.-Ing. Jürgen Stenzel
Erstreferent: Prof. Dr.-Ing. Ralf Steinmetz
Korreferent: Prof. Dr.-Ing. Ulrik Schroeder

Tag der Einreichung: 12. April 2011

Tag der Disputation: 17. Juni 2011

Darmstadt 2011
Hochschulkennziffer D17



Zusammenfassung

Die mit der Wissensgesellschaft verbundene Wissensexplosion, die veränderten Lebensbedingungen aufgrund neuer Arbeitsformen und die zahlreichen technischen Entwicklungen bedingen, dass das in Bildungseinrichtungen erworbene Wissen nicht mehr ein Leben lang ausreicht. Daher wird selbstgesteuertes Lernen am Arbeitsplatz immer wichtiger. Dies ist eine Form des Lernens, bei welcher ein aktueller Informationsbedarf durch die selbstständige Interaktion mit einer Vielzahl verschiedener digitaler Ressourcen gedeckt wird. Auf Grund dieser Interaktion wird das Lernen als Ressourcen-basiertes Lernen bezeichnet. Zunehmend wächst hierbei die Bedeutung des Web als Informationsquelle, denn es bietet eine Vielzahl an Ressourcen, die beim Lernen verwendet werden können.

Mit dem selbstgesteuerten Ressourcen-basierten Lernen sind jedoch auch Herausforderungen für die Lernenden verbunden. Zum einen handelt es sich bei digitalen Ressourcen im Web meist nicht um didaktisch aufbereitete Lernmaterialien. Die Ressourcen sind unstrukturiert und wurden nicht speziell für die Lernenden als Zielgruppe erstellt. Zusätzlich sind die relevanten Informationen oft über viele verschiedene Webseiten verteilt. Zum anderen gibt es bereits eine sehr große Menge an Informationen im Web, die sehr schnell zunimmt, was zur Informationsüberflutung führen kann. In dem in dieser Arbeit betrachteten Szenario des selbstgesteuerten Lernens gibt es zudem keinen Lehrenden, der den Lernprozess strukturiert. Die Lernenden müssen demzufolge selbstständig ihren Informationsbedarf feststellen und ihr Vorgehen planen. Sie müssen relevante Ressourcen identifizieren, annotieren und für eine spätere Nutzung organisieren. Dieser Umstand macht eine geeignete Verwaltung der Ressourcen notwendig. Die Mehrheit der Lernenden ist allerdings mit den aktuell vorhandenen Möglichkeiten für die Organisation von Webressourcen unzufrieden.

Ziel dieser Arbeit ist daher die Konzeption und Entwicklung eines Werkzeugs zur Unterstützung der Lernenden beim Ressourcen-basierten Lernen. Dies soll insbesondere die Verwaltung der Ressourcen durch die Lernenden unterstützen und die oben genannten Herausforderungen adressieren.

Das selbstgesteuerte Ressourcen-basierte Lernen verlangt ein persönliches Informations- und Wissensmanagement durch die Lernenden. In der Literatur existieren verschiedene Modelle zur Verwaltung von Informationen bzw. Wissen im organisationalen und persönlichen Bereich. Für selbstgesteuertes Ressourcen-basiertes Lernen fehlt ein solches Modell bisher. Daher wird zunächst ein Modell für das Ressourcen-basierte Lernen basierend auf den existierenden Modellen und einer im Rahmen dieser Arbeit durchgeführten Fragebogenerhebung entwickelt. Das Modell umfasst verschiedene Prozessschritte, die von dem zu entwickelnden Werkzeug unterstützt werden sollen.

Die Verwaltung der Ressourcen erfordert eine geeignete Speicherung dieser durch die Lernenden, beispielsweise themenbezogen oder aufgabenbezogen. Tagging stellt eine einfache und breit akzeptierte Möglichkeit zur Verwaltung von beliebigen Ressourcen im Web dar, allerdings ist die Ausdrucksmöglichkeit beim Tagging stark eingeschränkt. Andere Formen der Ressourcenverwaltung lassen sich im Bereich der formalen Wissensorganisation finden (z.B. Modellierung eines semantischen Netzes), allerdings wird zum Aufbau eines semantischen Netzes Expertenwissen vorausgesetzt. Als Basis für das im Rahmen der Arbeit entwickelte Werkzeug zur Unterstützung des Ressourcen-basierten Lernens wird daher eine Kombination beider Formen vorgeschlagen, d.h. ein semantisches Netz, das durch die Lernenden mittels

Tagging erstellt und erweitert wird. Kernbestandteil dieses Netzes sind Ressourcen und Tags. Beim Tagging wird jedem Tag ein Typ zugeordnet. So kann gespeichert werden, ob es sich z.B. um ein Thema oder eine Aufgabe handelt. Im Rahmen dieser Arbeit wird analysiert und evaluiert, welche Tag-Typen für das Szenario des Ressourcen-basierten Lernens notwendig sind.

Des Weiteren wird ein Algorithmus zur automatischen Erkennung dieser Tag-Typen vorgestellt, denn ein solcher Algorithmus kann den manuellen Pflegeaufwand bei der Verwaltung der Ressourcen reduzieren. Die Evaluation mit verschiedenen Korpora zeigt, dass dieser wissensbasierte Algorithmus ein Tag bereits während des Tagging-Vorgangs mit einer für das Szenario ausreichenden Genauigkeit klassifizieren kann.

Auf Basis des entwickelten Modells für das Ressourcen-basierte Lernen und dessen Anforderungen an die Verwaltung der Ressourcen werden verschiedene Werkzeuge und Systeme hinsichtlich ihrer Unterstützung des Ressourcen-basierten Lernens analysiert. Keines der verwandten Werkzeuge kann die Anforderungen angemessen erfüllen.

Daher wird ausgehend von den Prozessschritten im Modell und den daraus abgeleiteten funktionalen Anforderungen ein Konzept für ein Unterstützungswerkzeug entwickelt. Anhand der technischen Anforderungen wird ein System entworfen, bestehend aus Browser-Add-on, Backend zur Verwaltung der Wissensnetze und webbasiertem Frontend. Dieses wird implementiert und schließlich in Nutzerstudien evaluiert.

Die im Rahmen dieser Arbeit durchgeführten Nutzerstudien zeigen, dass die erweiterte Form des Taggings, basierend auf Tag-Typen, gut angenommen wird und eine angemessene Verwaltung der Ressourcen ermöglicht. Weiterhin zeigen die Studien, dass das implementierte Unterstützungswerkzeug die Herausforderungen des selbstgesteuerten Ressourcen-basierten Lernens angemessen adressiert. Die vorliegende Arbeit schafft damit eine Grundlage für die Optimierung des Vorgehens bei der selbstständigen Interaktion mit Ressourcen zur Deckung eines Informationsbedarfs.

Abstract

The knowledge explosion, changing circumstances due to new forms of work and many technical developments determine that the knowledge acquired in education is not sufficient throughout life. Therefore, self-directed learning in the workplace is becoming increasingly important. This is a form of learning where a current information need is met by the self-directed interaction with a wide range of digital resources. Therefore, this learning is called Resource-based Learning. Increasingly, the importance of the Web as an information source grows because it provides many resources that can be used for learning purposes.

However, self-directed Resource-based Learning also poses many challenges to learners. First, digital resources on the Web are usually not didactically prepared and therefore are not intended to be used as learning materials. In addition, the relevant information is often distributed across many different websites. Further, there is already a very large but still rapidly increasing amount of information available on the Web, which can lead to information overload. In the scenario of self-directed learning considered here, there is no teacher who structures the learning process. Therefore, learners have to independently determine their information needs and plan their proceeding. They have to identify, annotate and organize relevant resources for future use. This makes an appropriate management of resources necessary. However, the majority of learners is unsatisfied with the currently available possibilities for the organization of Web resources.

The goal of this thesis is therefore the design and development of a tool to support learners in Resource-based Learning. In particular, the management of resources should be supported and hence challenges mentioned above are addressed.

Self-directed Resource-based Learning requires a personal information and knowledge management by the learners. In literature, several models for managing information and knowledge in organizational and personal scopes exist. For self-directed Resource-based Learning such a model is missing so far. Therefore, a model for Resource-based Learning is developed based on the existing models and on a questionnaire survey conducted in the context of this thesis. This model encompasses several process steps that should be supported by the tool.

The management of resources necessitates the learners to appropriately store the resources, such as based on topic of interest or task to be executed. Tagging is a simple and accepted way to manage any resource on the Web, but its power of expression is restricted. Other forms of resource management can be found in the area of formal knowledge organization (e.g. modeling of a semantic network), however, expert knowledge is usually required to build a semantic network. As a basis for the tool that is developed in the context of this work, therefore, a combination of both forms is proposed, i.e. a semantic network that is created and expanded by the learners using tagging. Core components of this network are resources and tags. Additionally, a learner is able to assign a type to each tag. Therefore, the information whether the tag is e.g. a topic or task can be stored. As part of this thesis, the types of tags that are necessary for the scenario of Resource-based Learning have been analyzed and evaluated.

Furthermore, an algorithm for automatic detection of these tag types is presented, as such an algorithm can reduce the manual maintenance effort for the management of resources. The evaluation of various

corpora shows that the knowledge-based algorithm can classify a tag already during the tagging process with an accuracy which is sufficient for the scenario.

Based on the developed model of Resource-based Learning and its requirements for the management of resources, different tools and systems are analyzed with regard to their support of Resource-based Learning. None of the related tools fulfill the requirements appropriately.

Therefore, on the basis of the model's process steps and the derived functional requirements a concept for a supporting tool is developed. Based on the technical requirements, a system is designed, consisting of a browser add-on, a backend for the management of the knowledge networks and a web-based frontend. The tool is implemented and evaluated in user studies eventually.

The user studies conducted in this work show that the extended form of tagging, based on tag types, is well accepted and allows for appropriate management of resources. Furthermore, the studies show that the implemented tool addresses the challenges of self-directed Resource-based Learning adequately. The present work thus creates a basis for optimizing the approach to self-directed interaction with resources in order to meet an information need.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Motivation	1
1.2	Ziel, Ansatz und Beiträge der Arbeit	1
1.3	Aufbau der Arbeit	2
2	Ressourcen-basiertes Lernen	5
2.1	Exemplarische Lernszenarien	5
2.2	Grundlegende Definitionen	6
2.2.1	Ressourcen-basiertes Lernen	6
2.2.2	Lernressourcen	8
2.3	Analyse des Ressourcenmanagements beim Ressourcen-basierten Lernen	9
2.3.1	Ressource, Information und Wissen	9
2.3.2	Existierende Wissensmanagementmodelle	10
2.3.3	Ein Modell für Ressourcen-basiertes Lernen	18
2.4	Herausforderungen beim Ressourcen-basierten Lernen	23
2.4.1	Orientierungslosigkeit	24
2.4.2	Kognitive Mehrbelastung	24
2.5	Zusammenfassung	25
3	Bestehende Verfahren zur Wissensmodellierung	27
3.1	Wissensmodellierung durch Tagging	27
3.1.1	Vokabular	28
3.1.2	Akteure und Adressaten des Taggings	29
3.1.3	Funktionen von Tags	32
3.1.4	Tagging versus Klassifizierung	34
3.1.5	Zusammenfassung	35
3.2	Wissensrepräsentation	35
3.2.1	Taxonomien	37
3.2.2	Thesauri	37
3.2.3	Ontologien	37
3.2.4	Semantische Netze	38
3.2.5	Topic Maps	40
3.2.6	Folksonomien	40
3.3	Fazit	41
4	Typisiertes Tagging im Ressourcen-basierten Lernen	43
4.1	Semantisches Tagging	43
4.1.1	Definition	43

4.1.2	Semantische Tagging-Systeme	46
4.2	Facettierte Tagging	49
4.3	Typisiertes Tagging	52
4.3.1	Definition der Basismenge der Tag-Typen für Ressourcen-basiertes Lernen	52
4.3.2	Evaluation der definierten Basismenge	55
4.3.3	Nutzen der typisierten Tags	57
4.4	Recommendation von Tags und Ressourcen	57
4.4.1	Recommendation von Tags	58
4.4.2	Recommendation von Ressourcen	60
4.4.3	Fazit	60
4.5	Zusammenfassung	61
5	Automatische Erkennung der Tag-Typen	63
5.1	Motivation	63
5.2	Grundlagen und verwandte Ansätze	63
5.2.1	Eigennamenerkennung	64
5.2.2	Datenquellen für eine automatische Tag-Typ-Erkennung	64
5.2.3	Maße für den Vergleich von Klassifikationsverfahren	67
5.2.4	Existierende Ansätze zur Klassifikation von Tag-Typen	69
5.3	Ein wissensbasiertes Verfahren zur Erkennung von Tag-Typen	74
5.3.1	Verwendete Datenquellen und Vorgehen	75
5.3.2	Evaluation des eigenen Ansatzes	79
5.4	Zusammenfassung und zukünftige Arbeiten	86
6	Bestehende Ansätze und Werkzeuge zur Unterstützung des Ressourcen-basierten Lernens	89
6.1	Conceptual Diagram	89
6.2	Semantic Desktop	90
6.3	Personal Learning Environment	93
6.4	Social Bookmarking	95
6.5	Zielmanagement	98
6.6	Fazit	100
7	Basiskonzept und Design des Unterstützungswerkzeugs	103
7.1	Funktionale Anforderungen und Konzept	103
7.2	Technische Anforderungen und Design	111
7.2.1	Datenmodell	111
7.2.2	Anbindung an den Webbrowser	115
7.3	Zusammenfassung	115
8	Implementierung und Proof-of-Concept	117
8.1	Komponenten und Architektur	117
8.1.1	Backend: K-Infinity	117
8.1.2	Web-Service	119

8.1.3	Add-on	119
8.1.4	Gesamtarchitektur	119
8.2	Funktionen	122
8.2.1	Add-on	122
8.2.2	Knowledge-Portal	130
8.3	Zusammenfassung	131
9	Evaluation von ELWMS.KOM	133
9.1	Grundlegendes	133
9.1.1	Laboruntersuchung versus Felduntersuchung	133
9.1.2	Evaluationsmethoden	134
9.1.3	Herausforderungen bei Evaluationen im Kontext des Ressourcen-basierten Lernens	135
9.2	Durchgeführte Studien mit Vorläufern des Prototypen	136
9.2.1	Studie Labor 1	136
9.2.2	Studie Seminar 1	138
9.3	Durchgeführte Studien mit ELWMS.KOM	140
9.3.1	Studie Labor 2	140
9.3.2	Studie Seminar 2	142
9.3.3	Studie KOM	144
9.4	Zusammenfassung	150
10	Zusammenfassung und Ausblick	151
10.1	Zusammenfassung	151
10.2	Ausblick	152
	Literaturverzeichnis	153
	Abbildungsverzeichnis	167
	Tabellenverzeichnis	169
	Verzeichnis aller Abkürzungen	171
A	Details zum Verfahren der automatischen Erkennung der Tag-Typen	175
A.1	Korpus aus Eigennamen und englischsprachigen Tags	175
A.2	Berücksichtigung der Antwort der Nicht-Mehrheit	177
A.3	Erkennung von thematischen Zielen	178
A.4	Berücksichtigung der URL bei der Erkennung des Tag-Typs	179
A.5	DBpedia als weitere Datenquelle im Verfahren zur Erkennung des Tag-Typs	181
A.6	DBpedia als weitere Datenquelle und Berücksichtigung der URL im Verfahren zur Erkennung des Tag-Typs	183
A.7	Laufzeitverhalten des Verfahrens	186

B	Details zu durchgeführten Studien	187
B.1	[Studie Labor 2] - Auswertung des Fragebogens vor der Recherche	187
B.2	[Studie Labor 2] - Auswertung des Fragebogens nach der Recherche	190
B.3	[Studie Seminar 2] - Auswertung des Fragebogens nach der Nutzung von ELWMS.KOM .	193
B.4	[Studie KOM] - Fragebogen vor der Nutzung von ELWMS.KOM	196
B.5	[Studie KOM] - Auswertung des Fragebogens vor der Nutzung von ELWMS.KOM	205
B.6	[Studie KOM] - Fragebogen nach der Nutzung von ELWMS.KOM	207
B.7	[Studie KOM] - Auswertung des Fragebogens nach der Nutzung von ELWMS.KOM	215
C	Schriftenverzeichnis	217
C.1	Hauptveröffentlichungen	217
C.2	Mitautorenschaft und sonstige Veröffentlichungen	218
D	Lebenslauf des Verfassers	221
E	Erklärung laut §9 der Promotionsordnung	223

1 Einleitung

1.1 Motivation

Die mit der Wissensgesellschaft verbundene Wissensexplosion, die veränderten Lebensbedingungen aufgrund neuer Arbeitsformen und die zahlreichen technischen Entwicklungen bedingen, dass das in Bildungseinrichtungen erworbene Wissen nicht mehr ein Leben lang ausreicht. Daher wird selbstgesteuertes Lernen am Arbeitsplatz immer wichtiger. Zunehmend wächst die Bedeutung des Web als Informationsquelle, denn es bietet eine Vielzahl an Informationen, die beim Lernen verwendet werden können. In [91] wird dieses Lernen beschrieben als "ein Sich-verfügbarmachen von Informationen und Wissensbeständen bei aktuellen Problemen".

Diese Form des Lernens wird als Ressourcen-basiertes Lernen bezeichnet und ist das in dieser Arbeit betrachtete Anwendungsszenario. Streng genommen ist es keine neue Lernform, denn auch das Lernen mit einem Lehrbuch oder mit Studienbriefen kann zu Ressourcen-basiertem Lernen gezählt werden. Allerdings liegt der Fokus des Ressourcen-basierten Lernens im Sinne dieser Arbeit auf der Verwendung einer Vielzahl verschiedener Ressourcen insbesondere aus dem Web. Es wird in dieser Arbeit definiert als Lernen, bei welchem der aktuelle Informationsbedarf durch die selbstständige Interaktion mit einer Vielzahl verschiedener digitaler Lernressourcen gedeckt wird. Lernressourcen sind im Sinne dieser Arbeit alle digitalen Ressourcen, die das Potential haben, Lernen zu unterstützen.

Mit den Chancen des Ressourcen-basierten Lernens sind jedoch auch Herausforderungen für die Lernenden verbunden. Zum einen handelt es sich bei digitalen Ressourcen im Web meist nicht um didaktisch aufbereitete Lernmaterialien. Sie sind unstrukturiert und wurden nicht speziell für die Lernenden als Zielgruppe erstellt. Relevante Informationen sind oft über viele verschiedene Webseiten verteilt. Vielmals wird in der Literatur das Qualitätsgefälle der Informationen als weiteres Problem im Web genannt [73]. Zum anderen gibt es bereits eine sehr große Menge an Informationen im Web, die sehr schnell zunimmt, was zur Informationsüberflutung führen kann. In dem betrachteten Szenario gibt es zudem keinen Lehrenden, der Webseiten als Lernressourcen auswählt, sie in eine Reihenfolge bringt und auch das Vorwissen der Lernenden bei der Auswahl einbezieht. In der Hypertext-Forschung sind zudem Probleme wie Orientierungslosigkeit und kognitive Mehrbelastung bekannt, die durch die nicht-lineare Hypertext-Struktur auftreten und das Lernen mit Webressourcen behindern können.

Beim Ressourcen-basierten Lernen verwenden Lernende im Allgemeinen eine sehr große Anzahl von (Web-)Ressourcen. Dieser Umstand macht eine geeignete Verwaltung der Ressourcen notwendig. Die Mehrheit der Nutzer ist allerdings mit den aktuell vorhandenen Möglichkeiten für die Organisation von Webressourcen unzufrieden.

1.2 Ziel, Ansatz und Beiträge der Arbeit

Das Ressourcen-basierte Lernen bietet nicht nur Chancen, sondern auch Herausforderungen für die Lernenden. Ziel dieser Arbeit ist die Konzeption und Entwicklung eines Werkzeugs zur Unterstützung von Lernenden beim Ressourcen-basierten Lernen. Dies soll, da es ein zentrales Element ist, insbesondere

die Verwaltung der Ressourcen durch die Lernenden unterstützen und die zuvor genannten Herausforderungen adressieren.

Die vorliegende Arbeit enthält folgende Beiträge:

- Selbstgesteuertes Ressourcen-basiertes Lernen verlangt ein persönliches Informations- und Wissensmanagement durch den Lernenden. In der Literatur existieren verschiedene Modelle zur Verwaltung von Informationen bzw. Wissen im organisationalen und persönlichen Bereich. Für das selbstgesteuerte Ressourcen-basierte Lernen fehlt ein solches Modell bisher. Daher wurde ein Modell für das Ressourcen-basierte Lernen basierend auf den existierenden Modellen und einer in dieser Arbeit durchgeführten Fragebogenerhebung erstellt.
- Basierend auf diesem Modell und der durchgeführten Fragebogenerhebung wurden technische und funktionale Anforderungen für ein Unterstützungswerkzeug bestimmt und ein Konzept für dieses Werkzeug entworfen, das alle Prozessschritte des Ressourcen-basierten Lernens unterstützt.
- Tagging stellt eine einfache und breit akzeptierte Möglichkeit zur Verwaltung von beliebigen Ressourcen im Web dar, allerdings ist die Ausdrucksmöglichkeit beim Tagging stark eingeschränkt. Eine andere Form der Ressourcenverwaltung stellen formale Methoden zur Wissensorganisation dar (z.B. Modellierung eines semantischen Netzes), allerdings wird zum Aufbau eines semantischen Netzes Expertenwissen vorausgesetzt. Als Basis für das Werkzeug zur Unterstützung des Ressourcen-basierten Lernens wird daher eine Kombination beider Formen vorgeschlagen, d.h. ein semantisches Netz, das durch die Lernenden mittels Tagging erstellt und erweitert wird.
- Kernbestandteil des semantischen Netzes sind Tag-Typen, die durch Analyse der Literatur und durchgeführte Studien ermittelt wurden. Dies sind: Person, Ort, Ereignis, Typ (Genre), Thema und Ziel. Beim Tagging wird jedem Tag ein Typ zugeordnet.
- Es wurde ein Algorithmus zur automatischen Erkennung der Tag-Typen konzipiert, umgesetzt und auf unterschiedlichen Korpora evaluiert. Ein solcher Algorithmus kann den manuellen Pflegeaufwand bei der Verwaltung der Ressourcen reduzieren.
- Als Proof-of-Concept wurde das konzipierte Werkzeug umgesetzt. Es wurde eine Plattform entwickelt, bestehend aus Browser Add-on, Backend zur Verwaltung der Wissensnetze und webbasiertem Frontend.
- Schließlich wurde das Werkzeug in mehreren Benutzerstudien hinsichtlich seines Einsatzes im Ressourcen-basierten Lernen evaluiert.

1.3 Aufbau der Arbeit

Die vorliegende Arbeit gliedert sich wie folgt. Nach der Einleitung erfolgt in Kapitel 2 die Entwicklung eines Modells für das Ressourcen-basierte Lernen. Hierbei werden grundlegende Begriffe definiert, verwandte Modelle des Informations- und Wissensmanagements analysiert und beschrieben und Herausforderungen des Ressourcen-basierten Lernens erläutert. Eine der in dieser Arbeit adressierten Herausforderungen ist die Verwaltung der verwendeten Ressourcen durch die Lernenden. Daher werden in Kapitel 3 verschiedene existierende Verfahren zur Wissensmodellierung hinsichtlich eines Einsatzes zur Verwaltung der Ressourcen untersucht. Tagging wird in Kombination mit semantischen Netzen als geeignete Methode identifiziert, um die beim Lernen verwendeten Ressourcen zu verwalten. In Kapitel 4 wird

diese Kombination als typisiertes Tagging eingeführt. Beim typisierten Tagging wird jedem Tag ein Typ zugeordnet. So kann gespeichert werden, ob es sich z.B. um ein Thema oder eine Aufgabe handelt. Es wird analysiert und evaluiert, welche Tag-Typen für das Szenario des Ressourcen-basierten Lernens notwendig sind. Der Tag-Typ lässt sich auch automatisch bestimmen. In Kapitel 5 wird das in dieser Arbeit entwickelte und evaluierte Verfahren zur Erkennung von Tag-Typen vorgestellt.

Das typisierte Tagging bildet den Kern des in dieser Arbeit konzipierten Werkzeugs zur Unterstützung des Ressourcen-basierten Lernens. In Kapitel 6 werden daher verwandte Systeme und Werkzeuge beschrieben und hinsichtlich ihrer Unterstützung des Ressourcen-basierten Lernens analysiert. Anschließend wird in Kapitel 7 ein Konzept für das Unterstützungswerkzeug und in Kapitel 8 dessen Implementierung beschrieben. Die im Rahmen dieser Arbeit durchgeführten Benutzerstudien werden in Kapitel 9 beschrieben und diskutiert. Kapitel 10 fasst den Inhalt dieser Arbeit zusammen und schließt die vorliegende Arbeit mit einem Ausblick auf zukünftige Forschungsarbeiten ab.



2 Ressourcen-basiertes Lernen

Ressourcen-basiertes Lernen (RBL) ist das in dieser Arbeit adressierte Anwendungsszenario. Anhand ausgewählter Beispiele wird zu Beginn dieses Kapitels das Szenario des Ressourcen-basierten Lernens verdeutlicht. Im zweiten Abschnitt werden die grundlegenden Begrifflichkeiten definiert und erläutert. Ein zentrales Element des Ressourcen-basierten Lernens ist der Umgang mit Ressourcen und Wissen. Daher wird aufbauend auf Modellen des Wissensmanagements im dritten Abschnitt ein Modell für Ressourcen-basiertes Lernen entwickelt. Anschließend erfolgt im vierten Abschnitt eine Analyse der Herausforderungen für Lernende im adressierten Szenario. Ausgehend von dieser Analyse schließt das Kapitel mit der Formulierung von Anforderungen für ein System zur Unterstützung des Ressourcen-basierten Lernens.

2.1 Exemplarische Lernszenarien

Die folgenden Beispiele zeigen exemplarisch das Vorgehen von Lernenden bei der selbstständigen Recherche von Informationen, die zur Durchführung ihrer aktuellen Aufgabe oder zur Deckung ihres aktuellen Informationsbedarfs notwendig sind. Anhand der Beispiele wird verdeutlicht, welche Arten von Tätigkeiten durchgeführt, welche Informationsquellen und welche Werkzeuge genutzt werden. In den weiteren Abschnitten dieses Kapitels werden die Tätigkeiten, Informationsquellen und Werkzeuge genauer analysiert.

1. Beispiel - Vortragserstellung:

Birgit muss einen Vortrag zum Thema "Wind - ein Wetterphänomen und dessen Katastrophen" vorbereiten. Teil der Aufgabe ist die Präsentation des Themas in aktuellen Bildern und Zahlen. Für den Einstieg in das Thema erhält sie eine E-Mail vom Betreuer ihres Seminars mit einem Dokument über Grundlagen der Wetterphänomene im Anhang. Zusätzlich liest sie die Webseiten in Wikipedia¹ zum Thema Wind. In diesen Dokumenten findet sie eine Unterscheidung in verschiedene Arten von Wind, Informationen zu deren Entstehung und eine geographische Zuordnung. Mit diesen Informationen kann sie ihre Folienpräsentation grob strukturieren. Für die Details sucht sie im Web² nach weiteren Informationen. Sie speichert die Webseiten als Lesezeichen im Webbrowser und kopiert sich die relevanten Abschnitte der Webseiten in ein Dokument auf ihrer Festplatte. Bei ihren Recherchen findet sie neben allgemeinen Inhalten auch Nachrichten von kürzlich stattgefundenen Unwetterkatastrophen, Weblogs mit Augenzeugenberichten und wissenschaftliche Artikel mit aktuellen Forschungsergebnissen. Ein Teil dieser Webseiten ist relevant für Birgit, weil sie damit die Aktualität ihres Vortrages erhöhen und die Fotos gut als Bildmaterial verwenden kann. Sie speichert die Fotos in einem Ordner auf ihrer Festplatte. Die wissenschaftlichen Artikel pflegt sie in ihre Literaturdatenbank ein und notiert sich eine kurze Zusammenfassung. Später nutzt sie ihre Notizen und die gespeicherten Bilder, um ihre Folienpräsentation fertigzustellen.

¹ <http://www.wikipedia.org/>, Zuletzt abgerufen am 09.11.2010

² Web wird durchgängig verwendet als gebräuchliche Abkürzung für World Wide Web

2. Beispiel - Softwareprobleme:

In Philipps Firma wird für ein neues Projekt ein Team zusammengestellt. Sein Teamleiter beschließt, dass eine neue Software zur Projektdokumentation genutzt werden soll. Bei der Benutzung der Software kommt es gehäuft zu Programmabstürzen. Philipp durchsucht das Web nach Lösungen für sein Problem und findet verschiedene Foren und Listen mit häufig gestellten Fragen und Antworten (FAQ). Erst nach einigen Versuchen hat er den Grund gefunden und kann das Problem beheben. Neben der neuen Projektdokumentationssoftware muss er auch eine neue Programmiersprache erlernen. Da er bereits mehrere Programmiersprachen beherrscht, sucht er im Web nach Vergleichen zwischen den verschiedenen Sprachen, um deren Unterschiede kennenzulernen. Bei seinen Recherchen findet er außerdem nützliche Programmierbeispiele.

3. Beispiel - Reisevorbereitung:

Silvia und Tomas müssen dienstlich nach China fliegen, um mit potentiellen Projektpartnern zu sprechen. Sie werden China das erste Mal besuchen. Daher will sich Silvia vorher detailliert im Web über China informieren. Sie hat die Adresse ihres Hotels per E-Mail erhalten und beginnt mit der Suche nach Informationen zum Flughafen und Weg ins Hotel. Sie weiß, dass sich die chinesische Kultur von ihrer unterscheidet und sucht deshalb nach möglichen Fehlern, die sie vermeiden will. Sie kopiert die gefundenen Hinweise in ein Dokument auf ihrer Festplatte und erstellt eine sehr umfangreiche Liste. Sie findet einige englischsprachige Weblogs von Auswanderern, die über ihre Eindrücke in China berichten. Da Zeit für eigene Unternehmungen bleibt, sucht sie nach interessanten Ausflugszielen in der Nähe des Hotels. Tomas verbringt die Zeit damit, sein Englisch aufzufrischen und lernt vor allem Vokabeln, die speziell für seine Reise nach China wichtig sind.

2.2 Grundlegende Definitionen

2.2.1 Ressourcen-basiertes Lernen

Die Beispiele im vorangegangenen Abschnitt stellen das Vorgehen beim Ressourcen-basierten Lernen jeweils exemplarisch dar. In [91] wird diese Form des Lernens beschrieben als *"ein Sich-verfügbarmachen von Informationen und Wissensbeständen bei aktuellen Problemen"*. Eine Vielzahl weiterer Situationen ist vorstellbar. Die Beispiele decken das Anwendungsszenario des Ressourcen-basierten Lernens nicht vollständig ab, sondern dienen zur Einführung in die Thematik. In allen Beispielen wird deutlich, dass das Vorwissen der Personen nicht ausreicht, um den Anforderungen der aktuell gestellten Aufgabe gerecht zu werden. Die mit der Wissensgesellschaft verbundene Wissensexplosion, die veränderten Lebensbedingungen aufgrund neuer Arbeitsformen und die zahlreichen technischen Entwicklungen bedingen, dass das in Bildungseinrichtungen erworbene Wissen nicht mehr ein Leben lang ausreicht. In zunehmendem Maße erfolgt ein Einsatz von elektronischen Medien und Kommunikationstechnologien wie beispielsweise dem Internet. Diese Form des Lernens wird als E-Learning bezeichnet. E-Learning ist aus der heutigen Wissensgesellschaft nicht mehr wegzudenken, denn E-Learning ermöglicht den Zugriff auf Lernmaterialien mit der erforderlichen Flexibilität hinsichtlich Ort und Zeit und unterstützt somit berufs- und lebensbegleitendes Lernen. Nachteile der Bereitstellung elektronischer multimedialer Lernmaterialien sind allerdings der hohe Kostenfaktor bei ihrer Erstellung und didaktischen Aufbereitung [62, 93, 167]

und die für die Erstellung notwendige Zeit [83]. Daher sind nicht für alle denkbaren Aufgaben und Wissenslücken Lernmaterialien bzw. E-Learning-Kurse verfügbar. Zunehmend wächst die Bedeutung des Web als Informationsquelle, denn es bietet eine Vielzahl an Informationen, die im Ressourcen-basierten Lernen verwendet werden können.

Ressourcen-basiertes Lernen ist streng genommen keine neue Lernform, denn auch das Lernen mit einem Lehrbuch oder mit Studienbriefen kann zu Ressourcen-basiertem Lernen gezählt werden. Allerdings liegt der Fokus des Ressourcen-basierten Lernens im Sinne dieser Arbeit auf der zeitgleichen Verwendung einer Vielzahl verschiedener Ressourcen. Ressourcen-basiertes Lernen ist nicht mit einer bestimmten Lerntheorie verknüpft und kann beispielsweise seinen Einsatz im Rahmen von instruktionsorientierten, problembasierten, anfragebasierten oder projektbasierten Lernszenarien finden. Eine Übersicht von Studien zum Ressourcen-basierten Lernen findet sich in [57].

Eine frühe Definition von Ressourcen-basiertem Lernen, die elektronische Dokumente einschließt, gibt Breivik [19]: *"Resource-based learning occurs when teachers become facilitators of learning, helping students to learn from the information resources of the real world, such as books, journals, television, and online databases."* Demnach tritt Ressourcen-basiertes Lernen auf, wenn Lehrer zu Vermittlern des Lernens werden, die den Lernenden beim Lernen mittels Informationsressourcen der realen Welt (Bücher, Zeitschriften, Fernsehen und Online-Datenbanken) helfen. Diese Definition stammt aus der Zeit vor der Verbreitung des Internet. Mit Online-Datenbanken bezeichnet Breivik in ihren Arbeiten beispielsweise elektronische Bibliothekssysteme. Sie stellt in den Vordergrund, dass die Lehrenden u.a. in Zusammenarbeit mit Bibliotheken den Lernenden die Informationskompetenz vermitteln sollten, die im Rahmen des Ressourcen-basierten Lernens wichtig ist. Ihre Definition der Informationskompetenz basiert auf folgenden notwendigen Fähigkeiten [5]:

- wissen, wann ein Informationsbedarf besteht,
- identifizieren von Informationen, die nötig sind, um ein gegebenes Problem oder Thema zu adressieren,
- finden und bewerten von Informationen,
- organisieren von Informationen,
- Informationen effizient nutzen, um das Problem oder Thema zu adressieren.

Rakes [112] greift die Arbeiten von Breivik auf und definiert Ressourcen-basiertes Lernen wie folgt: *"Resource-based learning can be explained as a learning mode in which the student learns from his or her own interaction with a wide range of learning resources rather than from class exposition."* Sie versteht also Ressourcen-basiertes Lernen als Lernmodus, bei dem die Lernenden durch ihren eigenen Umgang mit einer breiten Palette von Lernressourcen anstatt durch Frontalunterricht lernen. Sie weist dem Internet als Lernressource eine zentrale Rolle im Ressourcen-basierten Lernen zu. Allerdings schränkt Rakes diese Rolle ein, indem sie vorschlägt, dass die Lernenden nicht selbst in der Menge aller Webseiten recherchieren, sondern in der Vorauswahl des Lehrenden, d.h. der Lehrende stellt im Vorhinein die möglichen Lernressourcen zusammen.

Tergan [149] greift die Definition von Rakes auf und bezieht sie vor allem auf Hypertextumgebungen und Internetressourcen. Die Lehrenden sollen Strategien für den Umgang mit Lernressourcen vermitteln und die Lernenden bei der Nutzung von Werkzeugen anleiten, die das Ressourcen-basierte Lernen unter-

stützen können. Im Gegensatz zu den beiden o.g. Arbeiten sollen die Lernenden unter Anleitung selbst nach relevanten Lernressourcen suchen.

Dies ist keine vollständige Nennung aller existierenden Definitionen des Ressourcen-basierten Lernens, weitere Definitionen sind zu finden in [57] und [123]. Einem Großteil der Definitionen gemein ist die Rolle des Lehrers. Der Lehrer vermittelt nicht den Lernstoff, wie sonst meist üblich mittels eines Vortrages oder vorbereitetem Lehrmaterial, sondern er leitet die Lernenden während der Durchführung des Ressourcen-basierten Lernens an und strukturiert den Lernprozess. Im lebensbegleitenden Lernen gibt es allerdings eher selten einen Lehrer, der konkrete Lernziele definiert und das Ressourcen-basierte Lernen anstößt; sondern der Impuls zur Durchführung Ressourcen-basierten Lernens entsteht durch einen aktuellen Informationsbedarf. Die Lernenden arbeiten selbstständig, dies wird auch als selbstgesteuertes Lernen bezeichnet.

Daher wird in dieser Arbeit Ressourcen-basiertes Lernen wie folgt definiert.

Definition 1 Ressourcen-basiertes Lernen wird definiert als Form des Lernens, bei welcher der aktuelle Informationsbedarf durch die selbstständige Interaktion mit einer Vielzahl verschiedener digitaler Lernressourcen gedeckt wird.

Diese Definition schließt die Anleitung durch einen Lehrer nicht prinzipiell aus, fokussiert aber auf das selbstgesteuerte Recherchieren und Lernen.

2.2.2 Lernressourcen

Im E-Learning gibt es verschiedene Definitionen für den Begriff *Lernressource*, die in der Literatur häufig auch als *Lernobjekt* bezeichnet wird. Das Learning Technology Standards Consortium (LTSC) des Institute of Electrical and Electronics Engineers (IEEE) definiert Lernobjekte als "*any entity, digital or non-digital, that may be used for learning, education or training*" [66]. Ein Lernobjekt ist demnach ein Objekt, gleichgültig ob digital oder nicht-digital, das zum Zweck des Lernens eingesetzt werden kann. Damit schließt diese Definition auch nicht-digitale Objekte ein. Unter nicht-digitale Objekte fallen beispielsweise auch Dinge wie Globus, Kompass, Wandkarte etc. Eine ähnliche Definition im Zusammenhang mit Ressourcen-basiertem Lernen wird in [60] gegeben, eine Ressource wird dort definiert als "*media, people, places, or ideas that have the potential to support learning*". Es werden also alle Medien, Personen, Orte oder Ideen, die das Potential haben Lernprozesse zu unterstützen, als Lernressourcen betrachtet.

Definitionen, die nicht-digitale Objekte oder Personen und Orte als Lernobjekte einbeziehen, sind für diese Arbeit zu breit. Die Definition "*A Learning Resource is a digital resource used for E-Learning*" in [119] schränkt die Definition einer Lernressource auf digitale Ressourcen, die im E-Learning verwendet werden, ein. Diese Definition setzt aber voraus, dass die Lernressource bereits im E-Learning eingesetzt wird, d.h. ursprünglich bereits für einen Einsatz im Lernkontext erstellt wurde. [159] erweitert daher die Definition einer Lernressource als "*any digital resource that can be reused to support learning*". Demnach muss eine Lernressource nicht zum Zwecke des Lernens erstellt worden sein, sondern fürs Lernen wiederverwendet werden können.

Die Begriffe Lernressource und Lernobjekt sind häufig mit gewissen Eigenschaften belegt, beispielsweise sollten Lernobjekte inhaltlich abgeschlossen, wiederverwendbar und leicht zugänglich sein. Die Zugänglichkeit kann beispielsweise durch die Beschreibung mit Metadaten erhöht werden, hierfür wird

häufig der im E-Learning etablierte Metadatenstandard Learning Object Metadata (LOM) [66] vorge schlagen. Des Weiteren wurden Formate wie das Shareable Content Object Reference Model (SCORM) [3] eingeführt, um die Interoperabilität zu gewährleisten. Diese Eigenschaften schränken die Menge der möglichen Lernressourcen für das Ressourcen-basierte Lernen zu stark ein.

Ausgehend von den o.g. Definitionen wird eine Lernressource in dieser Arbeit wie folgt definiert.

Definition 2 *Eine Lernressource wird definiert als digitale Ressource, die das Potential hat, Lernen zu unterstützen.*

Diese Definition schließt alle Definitionen ein, die sich auf digitale Ressourcen beschränken und erweitert die Menge der Lernressourcen dahingehend, dass die Lernressourcen nicht ursprünglich als Lernmaterialien erstellt oder didaktisch aufbereitet sein müssen. Sie müssen auch nicht die Eigenschaften wie inhaltliche Abgeschlossenheit, Zugänglichkeit durch Beschreibung mit Metadaten, Wiederverwendbarkeit oder bestimmte Formate einhalten. Damit umfasst diese Definition alle Webseiten, die nutzbare Informationen enthalten. Dazu zählen Wikis, Weblogs, Foren, Tutorials, Listen mit häufig gestellten Fragen und Antworten (FAQ), Nachrichtenseiten, Seiten über Hobbys, Firmen- und Produktbeschreibungen u.v.a. Auch Wissensdokumente, definiert in [82] als "*elektronisches Dokument, das Wissen enthält*", fallen unter diese Definition.

Weiterhin haben Lernressourcen, wie sie in dieser Arbeit betrachtet werden, folgende Eigenschaften:

- Sie sind primär textbasiert, können jedoch in Einzelfällen auch abgeschlossene multimediale Inhalte wie Bilder, Animationen oder Videos enthalten.
- Es handelt sich um Ressourcen, deren Zugriff über das Web möglich ist.
- Sie können eindeutig durch eine URL identifiziert werden.

Auf das Web als Informationsquelle für Ressourcen-basiertes Lernen wird in Abschnitt 2.4 näher eingegangen.

2.3 Analyse des Ressourcenmanagements beim Ressourcen-basierten Lernen

Im Ressourcen-basierten Lernen werden von den Lernenden viele Lernressourcen verwendet. Dieser Umstand macht eine geeignete Verwaltung der Ressourcen notwendig. In der Literatur existieren verschiedene Modelle zur Verwaltung von Informationen bzw. Wissen im geschäftlichen Bereich und für den persönlichen Bereich. In diesem Kapitel werden existierende Modelle analysiert und ihre Übertragbarkeit auf das in Abschnitt 2.2 definierte Szenario des Ressourcen-basierten Lernens untersucht.

2.3.1 Ressource, Information und Wissen

Für das weitere Verständnis ist es zunächst wichtig, zu klären, wie die Begriffe "Ressource", "Information" und "Wissen" in dieser Arbeit unterschieden werden. Das informationstechnische Verständnis unterscheidet Information und Wissen meist in einer stufenweisen Anordnung, die die Begriffe Zeichen, Daten, Information und Wissen enthält [10] und häufig als Wissenshierarchie oder Wissenspyramide bezeichnet wird. Zeichen werden durch Syntax zu Daten, Daten werden durch Hinzufügen von Bedeutung

zu Information, und Information, die in Handlungen umsetzbar ist, wird als Wissen definiert [42]. Abbildung 2.1 veranschaulicht die nach ihrem Sinngehalt gestufte Unterscheidung anhand eines Beispiels. Die Bildung eines Paares aus Uhrzeit und Dezimalzahl entspricht der Syntax, die aus den Zeichen auf der untersten Stufe Daten werden lässt. Die Interpretation dieser Daten als Uhrzeit und zugehörigem Stand des Börsenindex DAX fügt den Daten eine Bedeutung zu, damit werden die Daten zu Informationen. Die Informationen sind durch Hintergrundwissen umsetzbar in Handlungswissen, beispielsweise dass das Depot bei einem bestimmten DAX-Verlauf umgeschichtet werden sollte. Der Bedeutungsgehalt nimmt von der untersten Stufe zur obersten Stufe zu.

↑ steigender Bedeutungsgehalt	Wissen	"Depot muß umgeschichtet werden" "Der DAX ist ein Mittelwert"
	Information	(15:00 4487,0) (15:01 4487,1) DAX-Verlauf (15:02 4486,0)
	Daten	(15:00 4487,0)
	Zeichen	: \$ 0 ? 5 ß

Abbildung 2.1: Stufenweise Unterscheidung des Bedeutungsgehaltes [42]

Ähnliche Definitionen unterscheiden zusätzlich beispielsweise die Stufe der Weisheit (wisdom) oder treffen keine Aussagen zur Stufe der Zeichen. Einen Überblick über verschiedene Definitionen geben [120] und [78]. Die Informationswissenschaft weicht etwas von der üblichen Unterscheidung ab [78]: Daten sind im strengen Sinne gemessene bzw. beobachtete Einheiten (Messdaten) bzw. im weiteren Sinne nach syntaktischen Regeln gespeicherte Zeichen. Daten werden zu Informationen durch das gezielte Abrufen in einem bestimmten Kontext bzw. zu einem bestimmten Zweck. Durch die Umsetzung von Daten über Informationen in interne Wissensstrukturen - was Lernen entspricht - entsteht Wissen.

Eine Analyse aller existierenden Definitionen der Begriffe ist nicht Ziel dieser Arbeit, die hier genannten Definitionen sind ausreichend, um die Begriffe Information und Wissen anhand des Bedeutungsgehaltes zu unterscheiden und zu verstehen, wie aus Informationen Wissen werden kann. In dieser Arbeit wird neben den beiden Stufen noch die Stufe der Ressource unterschieden, die unter den beiden anderen angeordnet wird (siehe Abbildung 2.2). Sie entspricht den Daten nach [78].

2.3.2 Existierende Wissensmanagementmodelle

Ebenso wie für den Begriff des Wissens gibt es für den Begriff des Wissensmanagements verschiedene Definitionen, die unterschiedliche Aspekte wie Ziele, Methoden bzw. Prozesse beleuchten. Dabei wird Wissensmanagement aus verschiedenen Perspektiven betrachtet, zum einen aus der Perspektive der Organisation und zum anderen aus der Perspektive des Individuums.

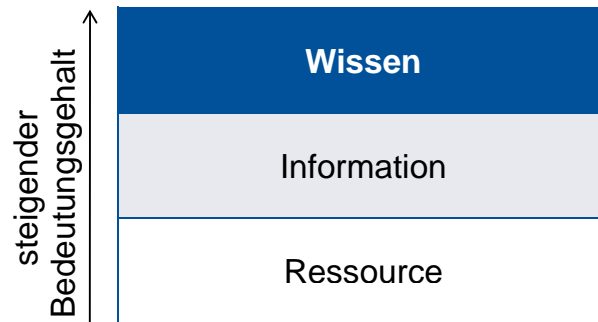


Abbildung 2.2: Stufenweise Unterscheidung von Ressource, Information und Wissen

Organisationales Wissensmanagement

Unter Wissensmanagement wird in [160] *"die Gesamtheit organisationaler Strategien zur Schaffung einer 'intelligenten' Organisation"* verstanden. Die Definition in [117] baut auf dieser auf und betrachtet die "intelligente" Organisation als lernende Organisation und bezeichnet den Kern des Wissensmanagement als *"systematischen und begründeten Umgang mit Wissen"*. Umgang umfasst dabei Rahmenbedingungen, Methoden, Werkzeuge und Prozesse. In [36] wird Wissensmanagement definiert als ein formaler Prozess zur Optimierung der Erzeugung, Verteilung und Nutzung von Wissen in einer Organisation, um das Wissen eines Unternehmens in Unternehmenswerte umzuwandeln. Die Autoren betrachten Wissen demzufolge als Unternehmenswert und fokussieren auf die Prozesse des Wissensmanagements, d.h. auf die Erzeugung, Verteilung und Nutzung von Wissen.

Für den Bereich des organisationalen Wissensmanagements werden häufig zwei Wissensmanagementmodelle zitiert (vgl. [118, 49, 158]), zum einen das sogenannte Modell der Wissensspirale [99] (begründet von Nonaka und Takeuchi) und zum anderen das sogenannte Modell der Bausteine des Wissensmanagements [108] (begründet von Probst et al.). Diese beiden Modelle beschreiben Wissensmanagement aus betriebswirtschaftlicher Sicht, d.h. wie Wissensmanagement in Organisationen initiiert oder optimiert werden kann, um den Erfolg eines Unternehmens zu steigern.

Das Modell der Wissensspirale (siehe Abbildung 2.3) beschreibt wie Wissen einer Person für andere Personen oder Personengruppen in einer Organisation nutzbar gemacht werden kann. Dabei wird angenommen, dass Wissen durch eine kontinuierliche, spiralförmige Wandlung zwischen implizitem und explizitem Wissen entsteht. Die vier in Unternehmen stattfindenden bzw. anzuregenden Transformationsprozesse sind definiert als Externalisierung (implizit zu explizit), Kombination (explizit zu explizit), Internalisierung (explizit zu implizit) und Sozialisierung (implizit zu implizit). Diese kontinuierlichen Transformationsprozesse wandeln im Unternehmen das personengebundene Wissen in allgemein zugängliches Wissen. Das Unternehmen muss hierzu die Wissensspirale durch geeignete Rahmenbedingungen in Gang setzen.

Das Modell der Bausteine des Wissensmanagement (siehe Abbildung 2.4) definiert sechs mögliche Kernprozesse im Umgang mit Wissen, die gegenseitig eng in Verbindung stehen und sich aufeinander auswirken: Wissensidentifikation, Wissenserwerb, Wissensentwicklung, Wissens(ver)teilung, Wissens-

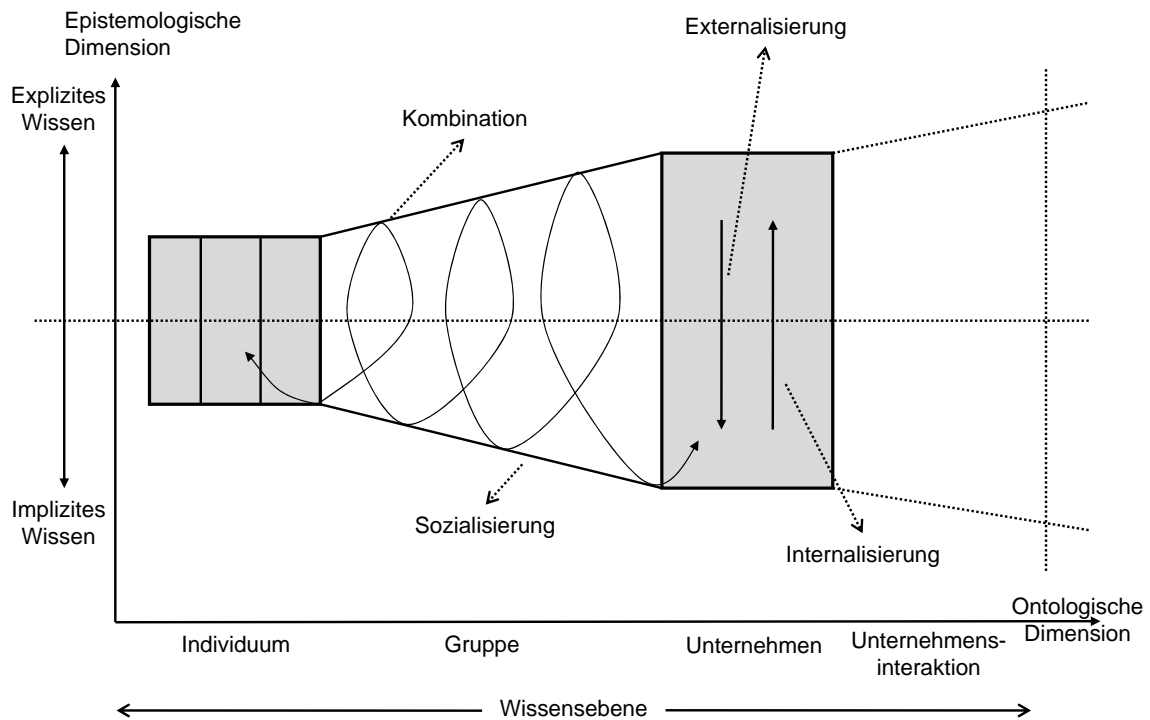


Abbildung 2.3: Modell der Wissensspirale (vgl. [99])

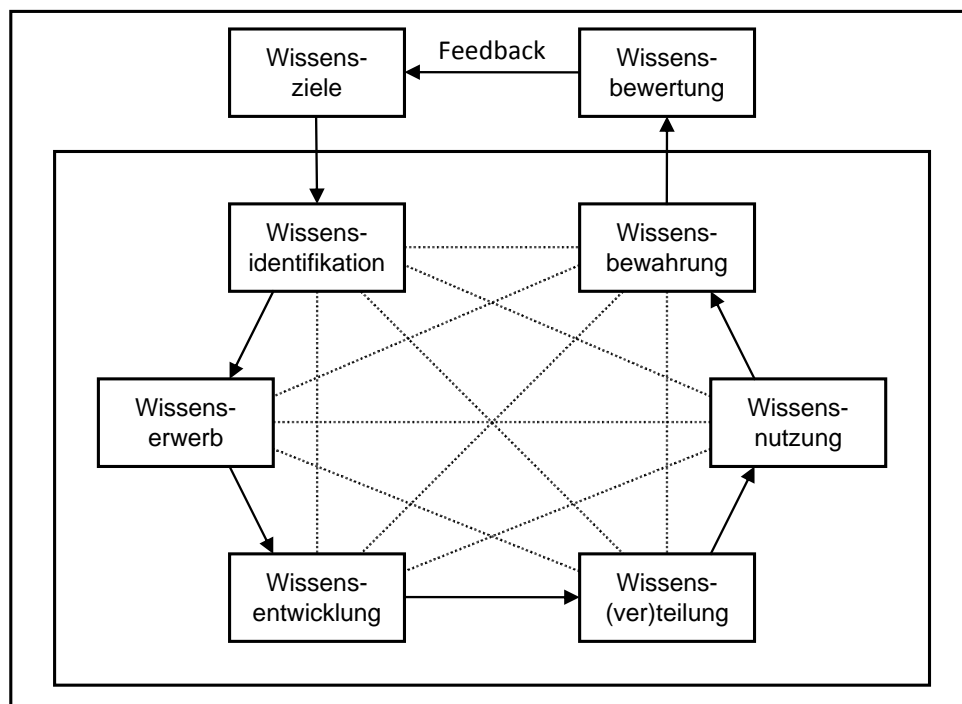


Abbildung 2.4: Modell der Bausteine des Wissensmanagements (vgl. [108])

nutzung und Wissensbewahrung. Die zwei weiteren Bausteine Wissensziele (Planung) und Wissensbewertung (Kontrolle der Umsetzung) ergänzen die Kernprozesse zu einem Managementregelkreis.

Neben diesen beiden weit verbreiteten Modellen gibt es in der Literatur eine Vielzahl weiterer Modelle.

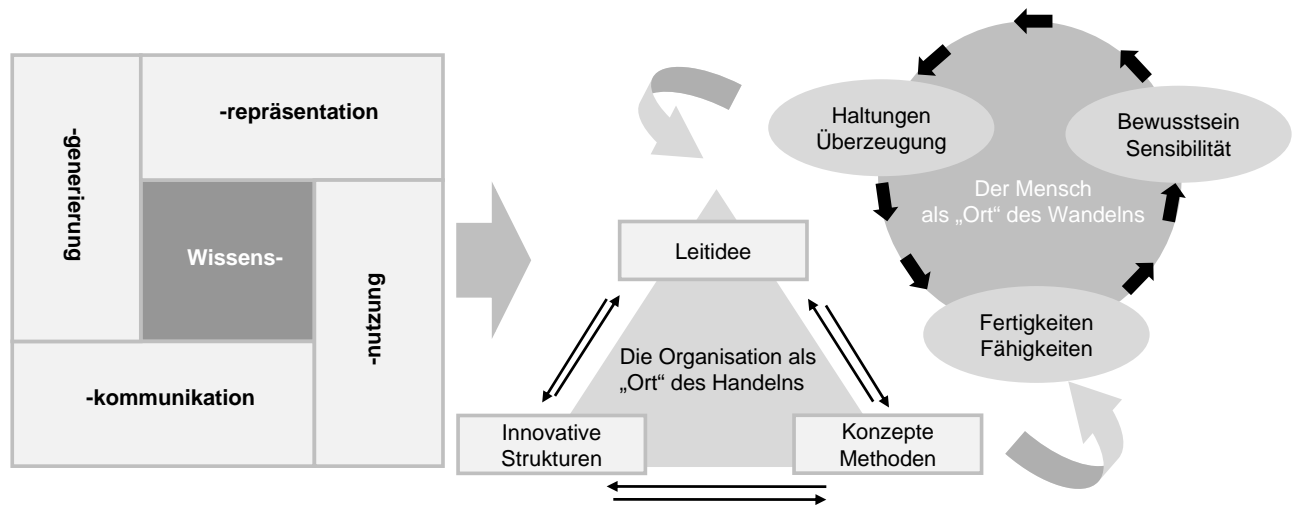


Abbildung 2.5: Münchener Modell (vgl. [117])

Ein Modell, das pädagogisch-psychologische Konzepte integriert, ist das Münchener Modell [118]. Dieses Modell (siehe Abbildung 2.5) definiert vier Kernprozesse in Anlehnung an das Modell der Bausteine (s.o.):

- Wissensrepräsentation,
- Wissensnutzung,
- Wissenskommunikation und
- Wissensgenerierung.

Voraussetzung und Ziel des Wissensmanagement soll die Anregung von organisationalen und individuellen Lernprozessen sein. Dabei müssen psychologische Voraussetzungen des Individuums wie Motive, Fähigkeiten und Kompetenzen mit einbezogen werden. Damit rückt dieses Modell das Individuum stärker in den Vordergrund. In den Mittelpunkt wird das Individuum beim persönlichen Wissensmanagement gesetzt.

Persönliches Wissensmanagement

In [59] wird die Kernaussage der Definitionen des persönlichen Wissensmanagements zusammengefasst als *"managing and supporting personal knowledge and information so that it is accessible, meaningful and valuable to the individual"*. Demnach ist persönliches Wissensmanagement die Verwaltung und Unterstützung von persönlichem Wissen, so dass es leicht zugänglich, sinnvoll und wertvoll für das Individuum ist. Im Gegensatz zu organisationalem Wissensmanagement wird die Verwaltung des persönlichen Wissens statt des organisationalen Wissens wichtig. Wissen wird wertvoll für das Individuum. Das Ziel des persönlichen Wissensmanagements wird in [29] beschrieben als *"empower each individual to easily apply their own personal knowledge in dealing with new and old problems, to learn from new experience and to create new knowledge"*. Das Individuum soll befähigt werden, persönliches Wissen beim Bearbeiten von

neuen oder alten Problemen anzuwenden, von den neuen Erfahrungen zu lernen und neues Wissen zu schaffen. Aufgabe des persönlichen Wissensmanagements ist das Bereitstellen eines Rahmens, durch den dieses Ziel erreicht werden kann.

Einige Definitionen betrachten die für persönliches Wissensmanagement notwendigen Fähigkeiten. Beispielsweise werden in [8] sieben Fähigkeiten definiert:

- Abrufen von Informationen,
- Bewertung von Informationen,
- Organisation von Informationen,
- Zusammenarbeit rund um Informationen,
- Analyse von Informationen,
- Darstellung von Informationen und
- Sicherung von Informationen.

Diese Fähigkeiten ähneln den im organisationalen Wissensmanagement definierten Kernprozessen, sind hier aber bezogen auf das Individuum.

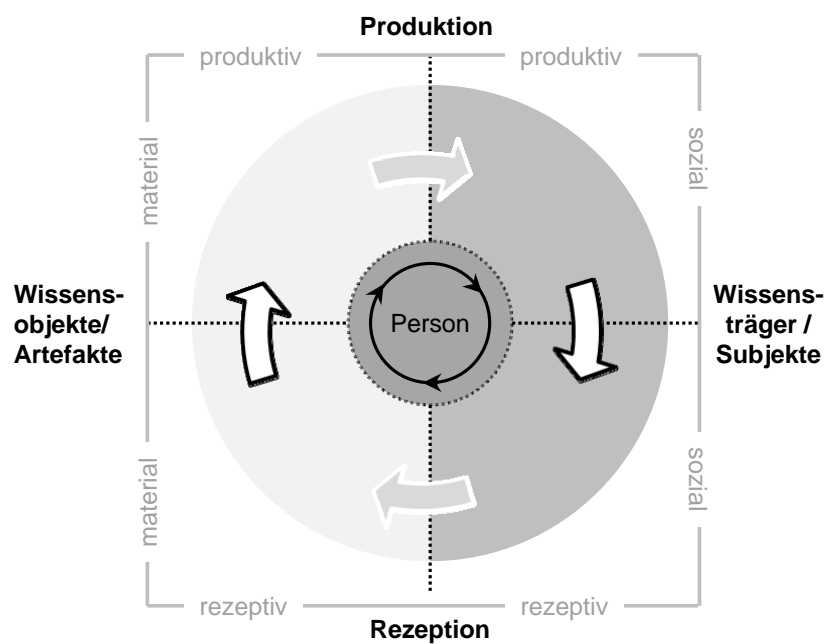


Abbildung 2.6: Modell zum persönlichen Wissensmanagement (vgl. [117])

Ein theoretisches Rahmenmodell für persönliches Wissensmanagement wird in [116] vorgeschlagen. Dieses Modell (siehe Abbildung 2.6) basiert auf folgenden grundlegenden Unterscheidungen. Es wird zwischen Innen und Außen des Wissensarbeiters unterschieden. Die Umwelt dieses Wissensarbeiters wird aufgeteilt in Personen und Gegenstände, mit denen der Wissensarbeiter interagiert. Des Weiteren wird die Aktivität unterschieden in Rezeption (z.B. Recherchieren, Lesen/Hören) und Produktion (z.B. Schreiben eines Artikels). Mentale Basisprinzipien (z.B. Metawissen, Lernstrategien und Problemlösekompetenz) bilden die Grundlage des persönlichen Wissensmanagements.

In der Literatur gibt es bisher wenig Modelle, die versuchen E-Learning und Wissensmanagement zu verbinden [117]. Eines dieser Modelle ist das Modell des Wissensreifungsprozess [125]. Das Modell beschreibt, welche individuellen Lernprozesse in Organisationen stattfinden können, durch die eine Art Wissensfluss entsteht, bei dem die Reife des Wissens entlang der Fließrichtung zunimmt. Abbildung 2.7 zeigt die fünf verschiedenen beobachteten Reifephasen. In der ersten Phase entstehen neue Ideen, die informell diskutiert werden. In der zweiten Phase bilden sich Communities, die auf Kollaborationsplattformen zusammen an den Ideen arbeiten und ein gemeinsames Vokabular formen. In der dritten Phase werden aus den bisher eher unstrukturierten Beiträgen beispielsweise Projektberichte oder andere strukturierte Dokumente produziert. Die Dokumente werden üblicherweise nicht als Lernmaterialien genutzt, sondern es werden in der vierten Phase Lernmaterialien erstellt und didaktisch aufbereitet, die in einer Ad-Hoc-Fortbildung genutzt werden können. In der fünften und höchsten Reifephase werden die Lernmaterialien aus der vierten Phase in Lernkursen zusammengestellt. Das Wissen ist in dieser Phase so weit gereift, dass es auch Anfängern bzw. Laien vermittelt werden kann, beispielsweise formal in Weiterbildungen. Zwischen den einzelnen Phasen gibt es Brüche, die den Wissensfluss behindern. Beispielsweise treffen eher informelle Lernphasen verknüpft mit Wissensmanagement (Phase 1-3) und eher formelle Lernphasen verknüpft mit klassischem E-Learning (Phase 4 und 5) aufeinander.

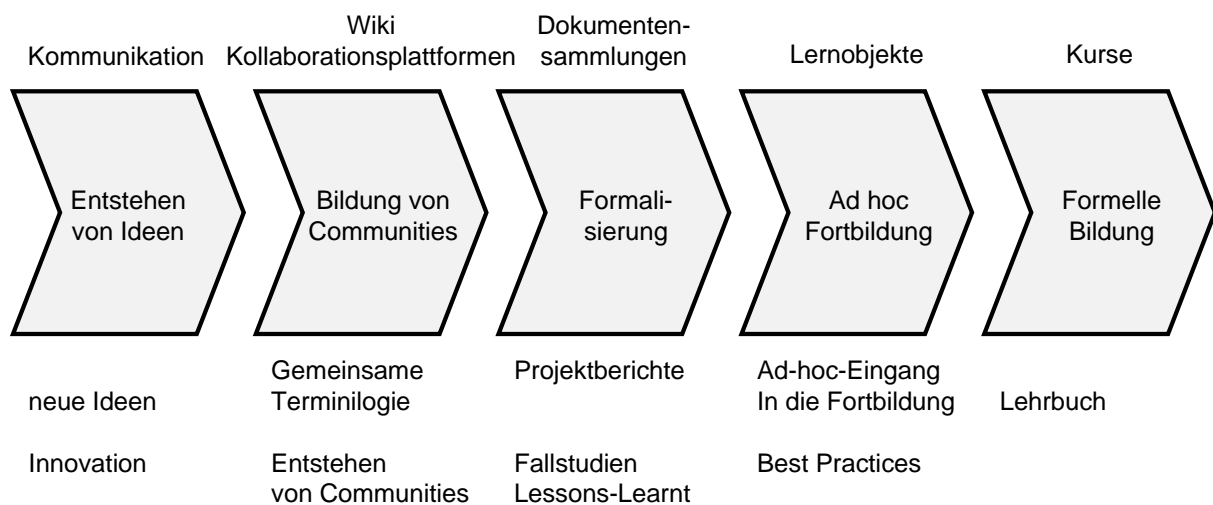


Abbildung 2.7: Wissensreifungsprozess (vgl. [125])

Dieses Modell fokussiert zwar in den einzelnen Phasen auf individuelle Lernprozesse, aber immer in Hinblick auf die Wissensweitergabe innerhalb einer Organisation oder eines Unternehmens. Im Modell wird das Reifen des Wissens durch Weitergabe von Ideen, Erstellen bzw. Ergänzen von Dokumenten und das Einbauen von neuem Wissen in vorhandene Wissensstrukturen als Lernen verstanden. Informelles Lernen jenseits der klassischen E-Learning-Schulungen kann somit als persönliches Wissensmanagement aufgefasst werden [117].

Tergan [149] greift das Modell der Bausteine des Wissensmanagements von Probst et al. auf (vgl. Abbildung 2.4) und modelliert persönliches Wissensmanagement für Ressourcen-basiertes Lernen. Er

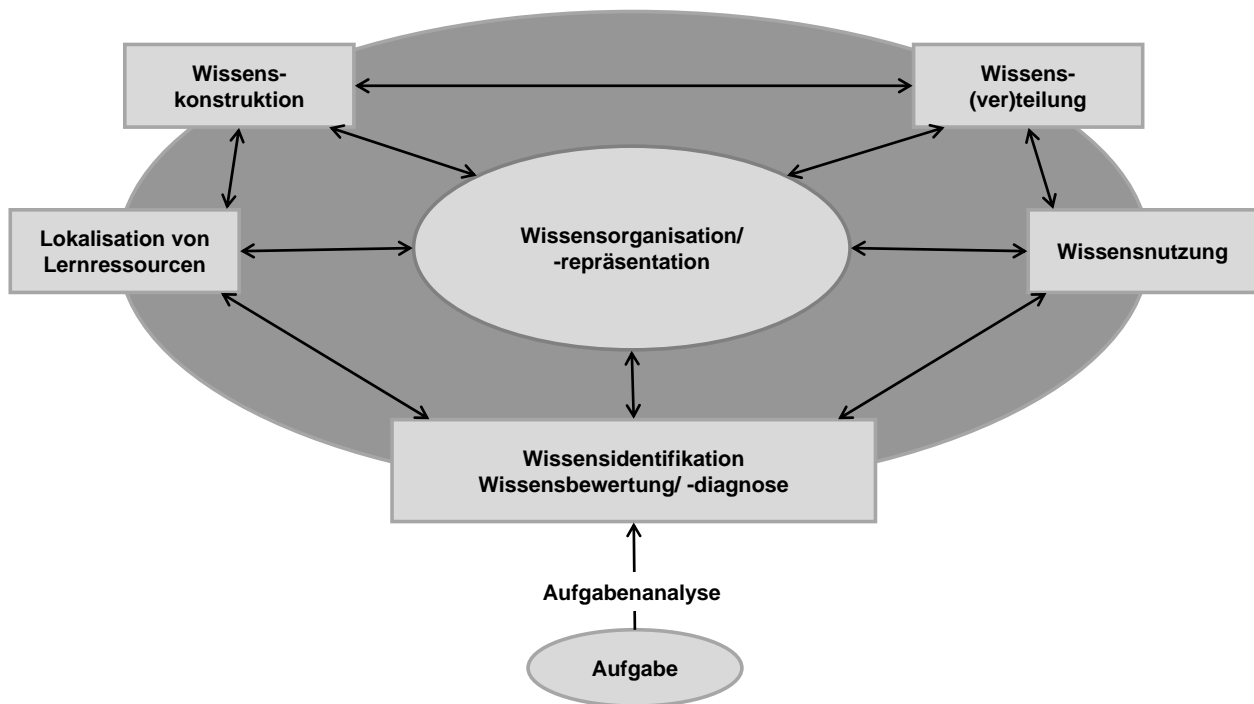


Abbildung 2.8: Bausteine des persönlichen Wissensmanagement (vgl. [149])

setzt den Baustein "Wissensorganisation/-repräsentation" an zentrale Stelle, weil seine Studien zeigen, dass dieser Baustein sehr wichtig ist.

- **Wissensidentifikation:** Am Anfang einer selbstgesteuerten Lernepisode steht eine Aufgabe, beispielsweise das Lösen eines Problems, eine bevorstehende Prüfung oder die Anfertigung einer Präsentation. Dieser Schritt dient dazu, sich einen Überblick zu verschaffen über bereits vorhandenes Wissen, das bei der Lösung der Aufgabe hilfreich sein kann, und festzustellen, welche Informationen zur Bearbeitung der Aufgabe noch fehlen.
- **Wissensbewertung / -diagnose:** Das bereits vorhandene Wissen, aber auch das in der Lernepisode erarbeitete Wissen wird in diesem Schritt in Bezug auf die zu bearbeitende Aufgabe hinsichtlich Quantität, Qualität und Funktionalität bewertet. Zum einen dient dieser Schritt als kontinuierliche Kontrolle der Aktivitäten, die parallel zu den anderen Wissensmanagementaktivitäten erfolgt. Zum anderen erfolgt in diesem Schritt während der Bearbeitung der Aufgabe oder am Ende eine explizite Kontrolle des Erfolges der Lernepisode, beispielsweise durch die Lösung des Problems oder durch weitere Testaufgaben.
- **Lokalisation von Lernressourcen:** Falls bei der Wissensidentifikation Wissenslücken festgestellt wurden, werden in diesem Schritt Lernressourcen gesucht, die für die Lösung der Aufgabe notwendig sind.
- **Wissenskonstruktion:** In diesem Schritt werden aktiv Wissensstrukturen aufgebaut. Dies wird auch als Wissenserwerb verstanden. Neues Wissen wird elaboriert und in bestehende Wissensstruk-

turen integriert, d.h. Informationen werden verarbeitet und mit vorhandenem Wissen verknüpft. Bestehende Wissensstrukturen werden umstrukturiert. Dieser Schritt ist eng verknüpft mit der Wissensrepräsentation und der Wissensnutzung.

- **Wissensorganisation / -repräsentation:** In diesem Schritt wird das konstruierte Wissen organisiert und geeignet repräsentiert. Dieser Schritt kann mental im Gedächtnis oder mittels externer Repräsentation (z.B. Papiernotizen, elektronisch) geschehen.
- **Wissensnutzung:** Wissen wird zur Bearbeitung der Aufgabe genutzt.
- **Wissens(ver)teilung:** Wissen wird miteinander kommuniziert, beispielsweise zur Wissensvermittlung in einem Lehr-/Lernkontext oder innerhalb einer Lerngruppe zur gemeinsamen Bearbeitung der Aufgabe. Die Wissens(ver)teilung kann dabei synchron (zeitgleich) oder asynchron (zu unterschiedlichen Zeitpunkten) erfolgen.

Fazit

Modelle des Wissensmanagements in Organisationen fokussieren zwar auf den Umgang mit Wissen in Unternehmen vor allem aus betriebswirtschaftlicher Sicht, aber sie enthalten auch Elemente, die für den persönlichen Umgang mit Wissen wichtig sind. Denn in Organisationen werden die Wissensmanagementaktivitäten von Individuen ausgeführt und deshalb spielen sie auch in organisationalen Wissensmanagementmodellen eine Rolle. Im Modell der Bausteine des Wissensmanagements von Probst et al. sind die Aktivitäten, die in den Prozessbausteinen stattfinden, teilweise auf den individuellen Prozess übertragbar, was im Modell der Bausteine des persönlichen Wissensmanagements von Tergan [149] gezeigt wurde. Das Modell der Wissensspirale erklärt die Entstehung von organisationalem Wissen durch Transformationsprozesse. Diese Transformationsprozesse werden durch Individuen ausgeführt und dehnen sich spiralförmig auf Gruppen von Individuen aus, was bereits in der Abbildung der ontologischen Ebene (Individuum-Gruppe-Unternehmen-Unternehmensinteraktion) zum Ausdruck kommt (siehe Abbildung 2.3). Das Münchener Modell reduziert die Anzahl der Bausteine auf vier Kernprozesse, allerdings sind die Kernprozesse sehr offen formuliert und es fehlen Prozesse wie Planung und Reflektion.

Modelle des persönlichen Wissensmanagements bauen häufig auf den Modellen des organisationalen Wissensmanagements auf und weisen dabei den Individuen die zentrale Rolle zu. Das Rahmenmodell zum persönlichen Wissensmanagement unterteilt die Aktivitäten in Rezeption und Produktion. Die Rezeption baut jedoch auf mentalen Prozessen auf, die sich schwierig von außen beobachten und analysieren lassen. Sie sind damit schwierig in Aktivitäten unterteilbar und in einem Modell integrierbar. Ziel des Modells der Wissensreifung ist die Optimierung des Prozesses der Erstellung von Lernmaterialien durch die Unterstützung des Wissensflusses im Unternehmen. Der Wissensfluss in Unternehmen und die Erstellung von Lernmaterialien ist nicht Gegenstand dieser Arbeit. Die Aktivitäten in den Phasen 1 bis 3 (Phasen des informellen Lernens) können auf das Ressourcen-basierte Lernen übertragen werden. Das Szenario des Ressourcen-basierten Lernens liegt dem Modell der Bausteine des persönlichen Wissensmanagements von Tergan bereits zu Grunde. Allerdings soll in dieser Arbeit analysiert werden, welche Tätigkeiten Lernende beim Ressourcen-basierten Lernen ausführen und welche Werkzeuge dabei

zum Einsatz kommen (können). Dafür ist das Modell von Tergan noch ungeeignet, weil zwar Ideen zur technischen Unterstützung von bestimmten Prozessen vorgeschlagen werden, aber die Tätigkeiten der Lernenden noch relativ allgemein formuliert sind.

2.3.3 Ein Modell für Ressourcen-basiertes Lernen

In diesem Abschnitt wird das in dieser Arbeit entwickelte Modell zur Beschreibung der Prozesse des Ressourcen-basierten Lernens vorgestellt. Das Modell basiert auf den in den vorhergehenden Abschnitten beschriebenen Wissensmanagementmodellen, im Besonderen auf dem Modell der Bausteine des persönlichen Wissensmanagements von Tergan [149] mit folgenden Unterschieden: Die Bausteine Wissenskonstruktion, Wissensrepräsentation und Wissensnutzung im Modell von Tergan sind abstrakt und überlappen sich teilweise. Diese Überlappung soll aufgelöst werden. Die Wissenskonstruktion ist kein eigener Prozessbaustein, sondern findet in verschiedenen Bausteinen statt. Dieser metakognitive Prozess lässt sich schwer beobachten. Er wird daher nicht als eigener Prozessschritt dargestellt.

Das Modell soll bei der Analyse der Anforderungen an und bei der Konzeption einer möglichen technischen Unterstützung des Ressourcen-basierten Lernens dienen. In einer Studie im Rahmen dieser Arbeit (siehe Kapitel 9 "Studie KOM") wurde daher auch erhoben, wie die Probanden beim Ressourcen-basierten Lernen vorgehen und welche Software bzw. Hilfsmittel sie dazu einsetzen.

Das vorgeschlagene Modell besteht aus den folgenden fünf Prozessschritten:

- **Planen & Reflektieren:**

Dieser Prozessschritt umfasst die Aufgabenanalyse, die Planung des Vorgehens und die Reflektion. Es muss identifiziert werden, welche Informationen zur Lösung der Aufgabe bereits bekannt sind und ob es Wissenslücken gibt. Wird das Vorgehen zur Lösung der Aufgabe zielgerichtet geplant, dann wird die Aufgabe dahingehend analysiert werden, ob und in welche Teilaufgaben sie zerlegt werden kann. Es werden sich Ziele überlegt, die während der Bearbeitung erreicht werden sollen. Gesetzte Ziele helfen bei der Beobachtung des Vorgehens und der Bewertung, ob die gesetzten Ziele erreicht wurden und die Aufgabe abgeschlossen werden kann. Während der Bearbeitung der Aufgabe muss eingeschätzt werden, ob die gerade verfolgte Strategie erfolgreich sein wird oder ob sie geändert werden muss. Die Reflektion dient auch dazu, zu analysieren, was bei einer zukünftigen Aufgabe besser gemacht werden kann.

38% der befragten Probanden planen das Vorgehen zur Bearbeitung einer Aufgabe und 85% passen ihre Strategie während einer Recherche an.

Die Bearbeitung der Aufgabe hängt neben der Aufgabe selbst auch von äußeren Faktoren ab. Dazu zählen beispielsweise die Zeit, die für die Bearbeitung der Aufgabe zur Verfügung steht, oder ein Termin, zu dem die Aufgabe gelöst sein muss. Häufig gibt es nicht nur eine einzige Aufgabe, sondern mehrere parallele Aufgaben. Einfache To-do-Listen können helfen, den Überblick über alle Aufgaben zu behalten. Eine andere Planungshilfe bieten Kalender, in denen belegte und freie Zeiten ersichtlich werden. Freie Zeiten können für die Bearbeitung der Aufgabe geblockt werden. Speziellere Planungssoftware bietet weitere Möglichkeiten, z.B. die Planung von Abhängigkeiten zwischen den einzelnen Aufgaben.

Wissenserwerb geschieht hier durch das Erlernen verschiedener Vorgehensweisen und das Reflektieren der Erfolge und Misserfolge insbesondere beim Vorgehen.

Die Bausteine Wissensidentifikation und Wissensbewertung/-diagnose aus dem Modell von Tergan können hier eingeordnet werden, führen aber nicht weit genug. Die Aufgabenanalyse ist dort im Modell außerhalb des Rahmens der Bausteine angeordnet. Die Aufgabenanalyse ist allerdings eng mit der Wissensidentifikation und Wissensbewertung verbunden, deshalb wurde die Aufgabenanalyse hier zu diesen Schritten zugeordnet und in einem Prozessschritt zusammengefasst.

- **Suchen:**

Dieser Prozess wird durchgeführt, wenn eine Wissenslücke bzw. ein Informationsbedarf festgestellt wurde. Es muss dann entschieden werden, in welcher Informationsquelle relevante Lernressourcen gefunden werden können. Der Bedarf kann entweder durch Lernressourcen gedeckt werden, die bereits in der Vergangenheit recherchiert wurden oder durch neu zu recherchierende Lernressourcen. Die Suchanfrage muss passend zur Informationsquelle formuliert werden. Die Suchergebnisse müssen dann hinsichtlich Qualität und Relevanz bewertet werden. Neben Volltext- oder Stichwortsuchen gibt es auch die Möglichkeit im Web zu browsen, beispielsweise in manuell erstellten Linklisten. Vor dem Auswählen einer Ressource wird anhand der Links und eventuell verfügbaren kurzen Beschreibungen entschieden, welcher Link als erstes bzw. nächstes angeklickt wird. Wird dann die Ressource aufgerufen, muss bewertet werden, ob sie relevant für die Suchanfrage ist und ob die Qualitätsanforderungen erfüllt sind.

Abhängig von den Suchergebnissen muss entweder die Suchstrategie oder das gesamte Vorgehen angepasst werden. Werden z.B. keine passenden Lernressourcen gefunden, muss die Suchanfrage umformuliert werden.

In der Befragung sollten die Probanden angeben, welche Informationsquellen sie zu welchem Anteil nutzen. Im Durchschnitt werden Suchmaschinen zu einem durchschnittlichen Anteil von 63% genutzt, Wikipedia zu 14% und die Suchmaschine für wissenschaftliche Publikationen Google Scholar³ zu 11%. Der restliche Anteil fällt auf Sonstiges zu je 2-4%. Auf die Frage der Art des Informationsbedarfs antworteten 94% der Probanden, dass sie nach Hintergrundinformationen für Vorträge/Schriftliche Ausarbeitungen/Publikationen suchen, nach wissenschaftlichen Publikationen suchen 83% und nach Illustrationsmaterial 68%. 51% suchen nach aktuellen Informationen aus Politik, Wirtschaft und Gesellschaft, 55% nach Zeitschriften und 40% nach Lehrmaterialien.

Viele Suchdienste gewichten die Suchtreffer, z.B. sortiert Google die Ergebnisliste mit Algorithmen wie PageRank [21]. Dies soll die Nutzer unterstützen, indem potentiell relevante Ergebnisse am Anfang der Liste erscheinen. Bewertungssysteme, in denen Nutzer Webseiten bewerten und kommentieren können, können bei der Qualitätseinstufung helfen.

Wissenserwerb kann beispielsweise stattfinden, wenn bei einer Suchanfrage festgestellt wird, dass es verschiedene Bedeutungen für einen Suchbegriff gibt, oder beim Bewerten der Suchergebnisse. Dieser Prozessschritt ist vergleichbar mit dem Prozessbaustein der Lokalisation von Lernressourcen.

³ <http://scholar.google.de/>, Zuletzt abgerufen am 03.12.2010

- **Annotieren & Organisieren:**

Dieser Schritt umfasst alle Tätigkeiten, die das Speichern der Informationen betrifft. Darunter zählen das Speichern der Lernressource z.B. in einem Ordner auf der Festplatte, die Beschreibung der Ressourcen mit Zusatzinformationen, das Strukturieren der Ressourcen und Informationen beispielsweise nach Themen, das Zusammenstellen von Themen in Begriffsstrukturen, das Umstrukturieren, usw. Dabei wird eine Vielzahl verschiedener Werkzeuge, meist parallel, eingesetzt.

Werkzeug	Probanden, die das Werkzeug verwenden
Ordner auf Festplatte	89%
E-Mail	60%
Texteditoren	85%
Tabellen	47%
Notizen auf Papier	72%
Literaturverwaltungssoftware	57%
Online-Bookmarking-Dienste	9%
Mind-Mapping-Werkzeuge	28%

Tabelle 2.1: Werkzeuge/Hilfsmittel zum Speichern von Rechercheergebnissen

Tabelle 2.1 listet die Antworten der von den Probanden üblicherweise eingesetzten Werkzeuge neben der Funktionalität des Webbrowsers auf. Am häufigsten werden die Ordner auf der Festplatte und Texteditoren zum Speichern genutzt. Das Papier zum Festhalten von Notizen wird gefolgt von der E-Mail und der Literaturverwaltungssoftware. Mind-Mapping-Werkzeuge werden immerhin noch von 28% der Probanden genutzt.

36% der Probanden gaben an, dass sie die Rechercheergebnisse direkt nutzen, d.h. 64% der Befragten organisieren ihre Rechercheergebnisse vor der eigentlichen Nutzung. Dabei sind knapp 50% mit dem Aufwand des Organisierens in Hinblick auf das Ergebnis eher unzufrieden.

Dieser Prozess hat große Auswirkung auf die anderen Prozesse. Je besser die Informationen annotiert und organisiert werden, desto besser können sie im Prozess des Suchens wieder aufgefunden werden. Je besser die Informationen in diesem Schritt aufgearbeitet werden, desto besser lassen sie sich später nutzen oder auch weitergeben.

Wissenserwerb kann hier beim Annotieren stattfinden, z.B. beim Schreiben einer Zusammenfassung. Das Organisieren anhand von Themen ermöglicht den Aufbau von Wissensstrukturen, was dem Wissenserwerb dienlich sein kann.

Dieser Prozessschritt entspricht der Wissensorganisation/-repräsentation im Modell von Tergan. Die Wissensrepräsentation umfasst Methoden und Werkzeuge, um Wissen in eine Form zu bringen, die von Computern interpretiert werden kann, um das modellierte Wissen zu verarbeiten und logisch zu schließen (*reasoning*) [145]. Auf die Wissensrepräsentation wird in Abschnitt 3.2 näher eingegangen.

- **Nutzen**

Dieser Prozessschritt umfasst alle Aktionen rund um die Nutzung der Informationen. Beispielsweise kann eine Aufgabe gelöst werden durch die Anwendung der richtigen Informationen, z.B. das Nachvollziehen von Codebeispielen beim Erlernen einer Programmiersprache. Informationen können bei der Erstellung von Präsentationen, Berichten, Artikeln etc. genutzt werden. Auch das Lesen und Lernen der Informationen ist diesem Schritt zuzuordnen.

Das tiefergehende Verstehen und Verarbeiten der Informationen in diesem Schritt ist vergleichbar mit der Wissenskonstruktion im Modell von Tergan. Dazu zählt auch das Kombinieren von Informationen, z.B. bei der Erstellung von Folien für eine Präsentation. Des Weiteren kann der Baustein Wissensnutzung aus dem Modell von Tergan hier eingeordnet werden, im Sinne des Nutzens von Informationen zur Lösung der Aufgabe.

- **Weitergeben:**

Dieser Prozess umfasst das Verteilen von Rechercheergebnissen und Informationen. Eine Weitergabe findet statt, weil eine Aufgabe gemeinsam bearbeitet wird oder um anderen zu helfen, beispielsweise durch das Weitergeben von früheren Ergebnissen. Ein Motiv für die Weitergabe kann auch die Selbstdarstellung sein.

Nur 10% der Befragten nutzen bisher Online-Bookmarking-Dienste zum Speichern von Rechercheergebnissen. Solche Dienste eignen sich gut zum Weitergeben von Informationen. 34% der Probanden gaben an, dass sie bereits solche Dienste zum Finden von Informationen anderer Benutzer genutzt haben. 92% der Probanden würden gern von den Recherchen ihrer Kollegen profitieren. Allerdings finden 55% der Befragten den Austausch von Suchergebnissen eher umständlich.

Neben der direkten Weitergabe von Informationen, beispielsweise mittels E-Mails, eignen sich Weblogs und Wikis zur Weitergabe von Wissen. Zahlreiche Studien untersuchen deren Einsatz im E-Learning und Wissensmanagement, u.a. [146, 76, 27, 48, 133, 9, 39, 126].

45% der Probanden haben bereits einen eigenen Weblog-Beitrag erstellt und 72% eine Wiki-Seite bearbeitet.

Wissenserwerb kann in diesem Schritt stattfinden, wenn die Informationen speziell für eine Weitergabe strukturiert und aufbereitet werden. Dieser Prozessschritt entspricht der Wissens(ver)teilung im Modell von Tergan.

Jeder Prozess des Ressourcen-basierten Lernens wirkt sich auf die anderen Prozesse aus. Eine Reihenfolge, die den Wechsel zwischen den Prozessen vorgibt, existiert nicht, sondern jeder Prozess kann in jeden anderen Prozess übergehen. So können beispielsweise Ergebnisse einer Suche direkt genutzt oder weitergegeben werden, ohne die Ergebnisse zu organisieren. Oder wenn bereits genügend Wissen zur Bearbeitung der Aufgabe existiert, kann der Prozess der Suche übersprungen werden. Werden mit einer Suche keine Ergebnisse erzielt, muss die Strategie angepasst werden, d.h. es passiert ein Wechsel zurück zum Prozessschritt Planen & Reflektieren. Abbildung 2.9 stellt das in dieser Arbeit erstellte Modell des Ressourcen-basierten Lernens dar.

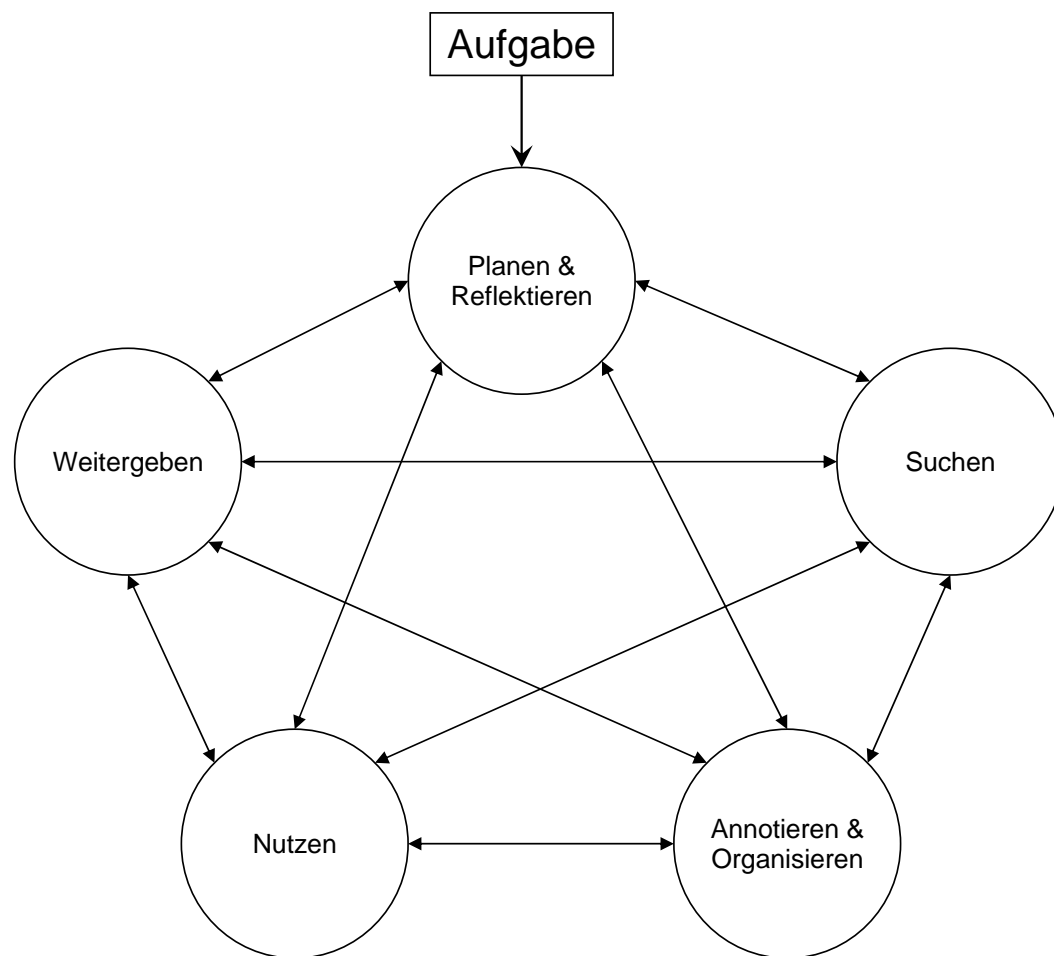


Abbildung 2.9: Modell der Prozessschritte für Ressourcen-basiertes Lernen

Persönliches Wissensmanagement und persönliches Informationsmanagement lassen sich oft nicht klar voneinander unterscheiden. In [124] wird persönliches Informationsmanagement definiert als "*management of data in the personal knowledge space as performed by the owning individual*". Es ist nach dieser Definition das individuelle Verwalten von Daten im persönlichen Wissensraum. Unter persönlichem Wissensraum versteht der Autor alle Daten, die vom Individuum während seiner Wissensarbeit genutzt oder erzeugt werden. Die Definitionen des persönlichen Wissensmanagements beziehen sich entweder auf die Verwaltung von Informationen (z.B. [8],[29]) oder die Verwaltung von Wissen (z.B. [154]) oder die Verwaltung von Informationen *und* Wissen (z.B. [51], [59]).

In dieser Arbeit werden die Begriffe Ressourcenmanagement, Informationsmanagement und Wissensmanagement auf Basis der Unterscheidung der Begriffe Ressource, Information und Wissen (siehe Abschnitt 2.3.1) unterschieden. Ressourcenmanagement meint die Ebene der Lernressourcen, also das Verwalten der Webseiten. Informationsmanagement bezieht sich beispielsweise auf Bibliographie-Angaben, die ein Lernender in eine Literaturverwaltungssoftware einträgt. Unter verwaltetem Wissen werden beispielsweise erstellte Notizen eines Lernenden mit Zusammenfassungen von Webseiten verstanden.

Gespeichertes Wissen eines Lernenden bedeutet Informationen für andere Lernende. Das Modell der Prozessschritte kann daher auf alle drei Ebenen angewendet werden, denn es können Ressourcen, Informationen und Wissen von einem Lernenden gesucht, organisiert, weitergegeben oder genutzt werden. In dieser Arbeit werden alle drei Ebenen aus Gründen der besseren Lesbarkeit im Begriff Ressourcenmanagement zusammengefasst, da dieser Begriff im Themenumfeld dieser Arbeit am wenigsten vorgebelegt ist. Soll eine Unterscheidung der Ebenen getroffen werden, wird dies explizit kenntlich gemacht.

2.4 Herausforderungen beim Ressourcen-basierten Lernen

Meist verlaufen die Aktivitäten beim Ressourcen-basierten Lernen nicht so geradlinig wie in den Beispielen aus Abschnitt 2.1 beschrieben:

Im Beispiel der Vortragserstellung hat Birgit ca. ein Jahr vorher einen Versicherungsbericht gelesen, der Versicherungsschäden der vergangenen zehn Jahre auflistete und beschrieb, wann, wo und in welcher Höhe Schäden durch Wind aufgetreten waren. Diesen Bericht hätte sie gut für ihre Präsentation nutzen können, um aktuelle Zahlen zu nennen, aber sie konnte diesen Bericht nicht wiederfinden. Manche der Webseiten versteht Anne nicht, da diese zu viel Vorwissen voraussetzen. Während ihrer Recherche muss sie leider eine längere Pause einlegen. Nach der Wiederaufnahme der Recherche kann sie sich nicht mehr erinnern, wo sie aufgehört hat und beginnt von Neuem. Bei manchen der gespeicherten Bildern hat sie vergessen, die Quelle zu notieren, so dass sie diese erneut heraussuchen muss, damit sie die Bilder in ihrer Präsentation verwenden kann.

Im Beispiel der Softwareprobleme treten bei einem Kollegen von Philipp ebenfalls Programmabstürze auf, aber Philipp kann sich nicht mehr genau an die Lösung erinnern. Leider hat er sich weder die Webseite gemerkt, die ihm damals weitergeholfen hat, noch hat er sich die Lösung notiert. Er sucht erneut im Web und findet die Webseite wieder, allerdings kann die Webseite aufgrund von Serverproblemen nicht geladen werden. Der gleiche Kollege hat sich vor ein paar Wochen bereits mit der neuen Programmiersprache beschäftigt und gute Links herausgesucht, allerdings weiß Philipp davon nichts.

Im Beispiel der Reisevorbereitung sind die Weblogs teilweise so interessant, dass Silvia sich durch weitere Einträge klickt und sich von den für die eigentliche Recherche irrelevanten Informationen ablenken lässt. Tomas stellt sich eine Liste mit wichtigen Vokabeln zusammen, damit er sie später lernen kann. Bei der Recherche nach Vokabeln, die besonders für China wichtig sind, stößt er auf viele Webseiten, die sich mit kulturellen Unterschieden und landestypischen Eigenarten beschäftigen. Silvia und Tomas hätten während ihrer Recherchen voneinander profitieren können, aber beide wussten nicht, dass sie parallel die gleichen Webseiten besuchten. Nach ihrer Reise hört ihr Vorgesetzter von ihren Listen und wünscht eine Verteilung an andere Mitarbeiter, leider fehlen die Quellen für diese Listen, so dass die Mitarbeiter nur umständlich Detailinformationen abrufen können.

In der Studie "KOM" (siehe Kapitel 9) gaben 85% der Probanden an oft und 13% manchmal für Studium oder Beruf zu recherchieren. 66% nutzen für ihre Recherchen immer das Web und die restlichen 32% sehr häufig.

Es gibt bereits eine sehr große Menge an Informationen im Web und die Informationsmenge nimmt zudem weiter zu, was zur Informationsflut führen kann.

45% der Probanden überfordert die große Menge an Informationen, die sie beim Recherchieren im Internet finden können. Oft wird in der Literatur das Qualitätsgefälle der Informationen als Problem im Web genannt, allerdings schätzen 85% der Probanden ein, dass sie die Qualität der Webseiten beurteilen können.

Die Webseiten sind meist keine didaktisch aufbereiteten Lernmaterialien. Sie sind unstrukturiert und nicht speziell für die Zielgruppe des Lernenden erstellt. Relevante Informationen sind über viele verschiedene Webseiten verteilt. Im Ressourcen-basierten Lernen, wie es in dieser Arbeit betrachtet wird, fehlt der Lehrer, der die Webseiten in eine richtige Reihenfolge bringt und auch das Vorwissen der Lernenden einbezieht. In der Hypertext-Forschung sind Probleme wie die Orientierungslosigkeit und die kognitive Mehrbelastung bekannt, die durch die nicht-lineare Hypertext-Struktur auftreten und das Lernen behindern können.

2.4.1 Orientierungslosigkeit

Eine Form der Orientierungslosigkeit wird in [34] als "getting lost in hyperspace" bezeichnet. Damit wird die Unsicherheit des Benutzers bei der Navigation im Hyperspace bezeichnet, wenn ein Benutzer nicht weiß, wo er sich aktuell in Bezug auf die gesamten Informationen befindet, oder wie er zu einer bestimmten Stelle im Hypertext gelangen kann. In [77] werden diese Formen der Unsicherheit durch eine Reihe weiterer ergänzt, die auf das Web als Hypertext-System übertragen werden können:

- Unsicherheit, den besten Startpunkt im Web zu finden,
- Schwierigkeiten, eine optimale Reihenfolge (Navigationspfad) zu finden, in welcher die Webseiten gelesen werden sollten,
- Schwierigkeiten beim Wiederauffinden bereits besuchter Webseiten und bei der Rekonstruktion vergangener Recherchen,
- Probleme, auf relevante Webseiten zurückzukehren nachdem mehrere irrelevanten Webseiten aufgerufen wurden (Rückwärtsnavigation),
- Unklarheit am Ende einer Recherche, ob alle wichtigen Informationen gefunden wurden und
- Unsicherheit beim Abschätzen der verbleibenden relevanten Informationen, die noch durchzusehen sind.

In [148] wird diese Art der Orientierungslosigkeit als strukturelle Desorientierung (*structural disorientation*) bezeichnet und von der konzeptuellen Desorientierung (*conceptual disorientation*) unterschieden. Konzeptuelle Desorientierung wird verursacht durch eine fehlende Integration neuer Informationen in die bereits existierende Wissensstruktur. Sie tritt beispielsweise auf, wenn relevantes Vorwissen fehlt, um die Informationen auf der Webseite zu verstehen. Der Zusammenhang von neuen Informationen und Vorwissen bleibt dann unklar. Diese Art von Desorientierung kann auch entstehen, wenn die semantische Beziehung zwischen Informationen verschiedener, aktuell gelesener Webseiten unklar bleibt. Der Aufbau einer zusammenhängenden Wissensstruktur wird dann schwierig.

2.4.2 Kognitive Mehrbelastung

Die kognitive Mehrbelastung (*cognitive overhead*) ist eng verknüpft mit der Orientierungslosigkeit. Sie ist jede zusätzliche Anstrengung und Konzentration über den eigentlichen Leseaufwand bzw. Lernaufwand

hinaus. Tätigkeiten, die beim Ressourcen-basierten Lernen zusätzlich zum reinen Lesen und Verstehen der Informationen kognitiven Aufwand erfordern, sind u.a.:

- Finden von Informationen
- Auswahl des als nächstes zu besuchenden Hyperlinks
- Bewertung der Informationen hinsichtlich Qualität und Relevanz
- Speichern der Informationen und der damit oft verbundenen Wechsel der Software-Applikation
- Thematische Einordnung der Informationen
- Verteilen der Informationen an andere Lernende

Eine zu hohe kognitive Mehrbelastung kann zur kognitiven Überlastung führen. Diese wiederum behindert Lernende bei der Wissensbildung erheblich (vgl. [96]). Wichtig wird deshalb die Unterstützung der Lernenden beim persönlichen Wissensmanagement, welches einen zentralen Bestandteil des Ressourcen-basierten Lernens darstellt, denn nur wenn der Aufwand für die verschiedenen Prozessschritte des Recherche- und Wissensmanagements gering genug ist, bleiben genügend kognitive Ressourcen frei für den Lernprozess. 84% der in der Studie "KOM" (siehe Kapitel 9) befragten Probanden finden es nützlich, eine Software zur Unterstützung ihrer Recherchen einzusetzen.

2.5 Zusammenfassung

Anhand von ausgewählten Beispielen wurde in diesem Kapitel das Szenario des Ressourcen-basierten Lernens vorgestellt und anschließend definiert. Ressourcen-basiertes Lernen wird in dieser Arbeit als eine Form des selbstgesteuerten Lernens verstanden, bei dem Lernende durch die Interaktion mit verschiedenen Lernressourcen lernen. Diese Lernressourcen müssen selbstständig von den Lernenden gesucht werden. Der Begriff der Lernressource wird weit gefasst; diese Arbeit fokussiert auf das Web als Informationsquelle und damit auf Webseiten als Lernressourcen. Diese Art des Lernens ist eng verknüpft mit persönlichem Wissensmanagement, daher wurden verschiedene Wissensmanagementmodelle vorgestellt, auf deren Basis ein Modell für das Ressourcen-basierte Lernen entwickelt wurde.

Das Ressourcen-basierte Lernen bietet aber nicht nur Chancen, sondern auch Herausforderungen für die Lernenden. Ziel dieser Arbeit ist die Entwicklung eines Werkzeugs zur Unterstützung von Lernenden beim Ressourcen-basierten Lernen.

Folgende Kriterien soll das Werkzeug erfüllen:

- Reduktion der strukturellen Desorientierung
- Reduktion der konzeptuellen Desorientierung
- Vermeidung der kognitiven Überlastung durch Reduktion der Mehrbelastung
- Unterstützung aller Prozessschritte des persönlichen Wissensmanagements
- Unterscheidung von Ressourcen-, Informations- und Wissensmanagement darf im Zielsystem für die Lernenden keine praktische Relevanz haben, sondern die Verwaltung aller drei Ebenen muss in einem einzelnen Werkzeug möglich sein.



3 Bestehende Verfahren zur Wissensmodellierung

Zur Unterstützung des Ressourcenmanagements im Ressourcen-basierten Lernen wird ein System benötigt, das das Speichern von Ressourcen, Informationen und Wissen ermöglicht. Dies entspricht dem Prozessschritt "Annotieren & Organisieren" im Modell des Ressourcen-basierten Lernens (siehe Kapitel 2.3.3) und ist Basis für die anderen Prozessschritte. Bezüglich der Organisation von Wissen spricht man in der Informatik und in den Informationswissenschaften von *Wissensmodellierung*. Ziel der Wissensmodellierung ist die Verwaltung von Wissen zum Zweck der Auffindbarkeit bzw. Wiederauffindbarkeit in Wissensbasierten Systemen. Mögliche Formen der Wissensmodellierung werden in Abschnitt 3.2 erläutert und auf einen Einsatz im Anwendungsszenario des Ressourcen-basierten Lernens hin untersucht. Die Pflege einer Wissensbasis ist im Allgemeinen sehr aufwendig und wird innerhalb von Organisationen in der Regel von Experten durchgeführt. Im Falle des Ressourcen-basierten Lernens muss der Aufbau und die Pflege der Wissensrepräsentation durch die Lernenden selbst durchgeführt werden. *Tagging* kann eine leichtgewichtige Form zur Erstellung einer Wissensrepräsentation darstellen. Im folgenden Abschnitt wird auf Tagging grundlegend eingegangen.

3.1 Wissensmodellierung durch Tagging

Tagging-Systeme stellen eine Möglichkeit zum Management und zur Organisation von Objekten dar [95]. Beim Tagging werden Schlagworte zu Objekten zugeordnet mit dem Ziel, das Finden bzw. Wiederfinden dieser Objekte zu ermöglichen. Tags sind somit Metadaten in Form von Schlagworten. Auf den Zusammenhang von Tags und Metadaten wird im Abschnitt 3.1.2 kurz eingegangen. Tagging-Systeme sind im Web stark verbreitet. So lassen sich digitale Objekte wie Bilder in Flickr¹, Videos in YouTube², Webseiten in Delicious³, E-Mails in GoogleMail⁴ und Kontakte in Xing⁵ taggen, um nur einige Beispiele aufzulisten.

In einem Tagging-System lassen sich die Begriffe *Tag*, *Tag-Instanz* und *Tag-Bezeichner* wie folgt unterscheiden (vgl. [95]). Eine Tag-Instanz ist ein Tupel der Form:

$$(\text{Objekt}, \text{Tag-Bezeichner}, \text{Benutzer})$$

Die Objekte müssen eindeutig referenzierbar sein. Objekte, wie sie in dieser Arbeit hauptsächlich betrachtet werden, sind Ressourcen wie sie in Kapitel 2.2 definiert wurden, d.h. es sind Webseiten, die durch ihre URL eindeutig referenzierbar sind.

Beim Tagging lassen sich zwei Fälle unterscheiden: entweder können die Tag-Bezeichner vom Benutzer frei vergeben werden oder sie sind vordefiniert und der Benutzer darf lediglich aus einer Menge vorgeschlagener Tag-Bezeichner auswählen. Im Abschnitt 3.1.1 wird auf die Vorteile und Nachteile eines freien Vokabulars eingegangen. Benutzer sind die Anwender, die das Tag bzw. die Tag-Instanz erstellen.

¹ <http://www.flickr.com/>, Zuletzt abgerufen am 02.01.2011

² <http://www.youtube.com/>, Zuletzt abgerufen am 02.01.2011

³ <http://www.delicious.com/>, Zuletzt abgerufen am 02.01.2011

⁴ <http://gmail.com/>, Zuletzt abgerufen am 02.01.2011

⁵ <https://www.xing.com/>, Zuletzt abgerufen am 02.01.2011

Werden zusätzlich noch Daten wie der Zeitpunkt des Taggings gespeichert, ergibt sich ein n-Tupel der Form:

$$(Objekt, Tag-Bezeichner, Benutzer, Zeitstempel, \dots)$$

Ein Tag setzt sich zusammen aus dem Tag-Bezeichner und der Menge aller Objekte, die mit diesem Tag-Bezeichner ausgezeichnet wurden:

$$(Tag-Bezeichner_i, \{Objekt_j \mid \exists (Objekt_j, Tag-Bezeichner_i, Benutzer_x)\})$$

In der Literatur und in Tagging-Systemen wird häufig der Begriff Tag auch für den Begriff des Tag-Bezeichners verwendet. So wird die Trennung dieser Begriffe auch in dieser Arbeit nur dann vorgenommen, wenn die Unterscheidung von Tag-Instanz und Tag-Bezeichner wichtig ist; im Allgemeinen wird jedoch der Begriff Tag verwendet.

Die Organisation von Objekten mittels Verschlagwortung ist nicht neu. Kategorisierung bzw. Indexierung wird beispielsweise im Bibliothekswesen durch Experten wie Bibliothekare durchgeführt [120]. Ebenso war die Organisation z.B. der persönlichen, digitalen Musiksammlung mittels Stichwörtern bereits vor der Etablierung des Taggings möglich. Im Unterschied zu diesen Ansätzen enthält das Tagging eine soziale Komponente, d.h. es ist möglich, Tags (und getaggte Objekte) mit anderen auszutauschen. Dieses gemeinschaftliche Tagging wird als *Social Tagging* bezeichnet. Im Abschnitt 3.1.2 werden die verschiedenen Akteure und Adressaten des Taggings detailliert beschrieben. Die verschiedenen Motive beim Tagging werden in Abschnitt 3.1.3 betrachtet. Im Gegensatz zur monohierarchischen Klassifikation können pro Objekt mehrere Tag-Instanzen erstellt werden, in Abschnitt 3.1.4 werden die Vorteile des Taggings gegenüber der monohierarchischen Klassifizierung herausgearbeitet.

3.1.1 Vokabular

Das Vokabular in Tagging-Systemen lässt sich unterscheiden in kontrolliertes und unkontrolliertes Vokabular. Hierbei wird als Tagging mit unkontrolliertem Vokabular der Prozess bezeichnet, bei dem die Benutzer in ihrer Wahl des Tag-Bezeichners nicht eingeschränkt sind. Unter kontrolliertem Vokabular wird ein vordefiniertes, fest vorgegebenes Vokabular verstanden. Potentiale und Nachteile der beiden Varianten werden u.a. diskutiert in [87], [47] und [18]. Dabei stellen die Potentiale des kontrollierten Vokabulars gleichzeitig die Nachteile des unkontrollierten Vokabulars dar, und umgekehrt.

Für kontrollierte Vokabulare gelten u.a. die folgenden Nachteile:

- Die Nutzung des Vokabulars erfordert Wissen über das Vokabular und erhöht die Komplexität beim Tagging.
- Das Vokabular muss im Vorhinein erstellt werden. Bei unzureichender Pflege kann es veralten. Aufgrund der großen Menge an verschiedenen Informationen im Web, zu denen ständig neue Informationen hinzukommen, ist die Definition eines vollständigen Vokabulars nicht durchführbar.
- Der Aufbau und die Pflege des Vokabulars kann hohe Kosten verursachen und erfordert Expertenwissen.

Ein unkontrolliertes Vokabular muss nicht vorab erstellt werden, ist demzufolge immer aktuell und die Vergabe von Tags erfordert kein Expertenwissen. Dafür besitzt es die folgenden Nachteile, die durch ein kontrolliertes Vokabular vermieden werden können:

- Häufigere Rechtschreibfehler
- Mehrsprachigkeit der Tag-Bezeichner
- Uneinheitliche Schreibweisen, z.B. e-Learning oder eLearning
- Sonderzeichen, z.B. dringend!!!, *interessant*
- Mischung aus Singular und Plural
- Tags auf verschiedenen Abstraktionsebenen, z.B. Perserkater, Katze, Tier
- Gebrauch von Synonymen, d.h. gleiche Objekte werden mit verschiedenen Tag-Bezeichnern versehen, z.B. Ferien und Urlaub
- Verwendung von Homonymen, d.h. verschiedene Objekte werden mit gleichen Tag-Bezeichnern versehen, z.B. Java für die Programmiersprache, für die Insel und für den Kaffee
- Subjektive Tags, die nur für den Einzelnen von Bedeutung sind

Diese Nachteile können dazu führen, dass Objekte nicht gefunden werden oder dass Tags einen irreführenden Zusammenhang zwischen Objekten herstellen, die durch gleichlautende Tag-Bezeichner verursacht werden.

3.1.2 Akteure und Adressaten des Taggings

Die Verschlagwortung von Objekten kann von verschiedenen *Akteuren* durchgeführt werden. Dabei kann es ein Ziel der Akteure sein, die Tags selbst zu nutzen oder die Objekte für andere Nutzer zu taggen; somit lassen sich die *Adressaten* der Tags in die Gruppe der Ersteller der Tags und in die Gruppe der Konsumenten unterscheiden. Die Akteure lassen sich weiter unterteilen in die Gruppe der Ersteller der Objekte und in die Gruppe der Konsumenten der Objekte.

Die beiden Abbildungen 3.1 und 3.2 veranschaulichen die Einteilung in die verschiedenen Akteure und Adressaten des Taggings. Der Abschnitt 3.1.3 geht näher auf die Motivation, für sich selbst oder andere Nutzer zu taggen, und auf die Funktionen von Tags ein.

In [56] werden die Akteure und die Adressaten jeweils eingeteilt in die Autoren und Konsumenten der Inhalte bzw. in die Ersteller und Konsumenten der Tags. Daher ergibt sich die in Abbildung 3.1 dargestellte Vier-Felder-Matrix. Für jedes Feld sind ein oder mehrere Tagging-Systeme angeführt, die anhand der Literatur als exemplarische Vertreter dieser Gruppe zugeordnet werden können. Diese Zuordnung ist nicht exklusiv, sondern es gibt auch einzelne Beispiele, in denen die Tagging-Systeme in andere Gruppen eingeordnet werden können.

Feld 1 - Selbst-Selbst: Diese Gruppe umfasst die Autoren, die ihre eigenen Inhalte zu ihrem eigenen Nutzen taggen. Ein Beispiel hierfür ist der Photo-Dienst Flickr, bei dem Nutzer Bilder hochladen und mit Tags organisieren können. Primäres Ziel des Taggings dieser Gruppe ist dabei zumeist die Ordnung der eigenen Bildersammlung, ohne andere Adressaten zu bedenken.

Feld 2 - Andere-Selbst: In diese Gruppe können die Konsumenten der Inhalte eingeordnet werden, die fremde Inhalte für die eigene Sammlung taggen, um die Inhalte später wiederfinden zu können.

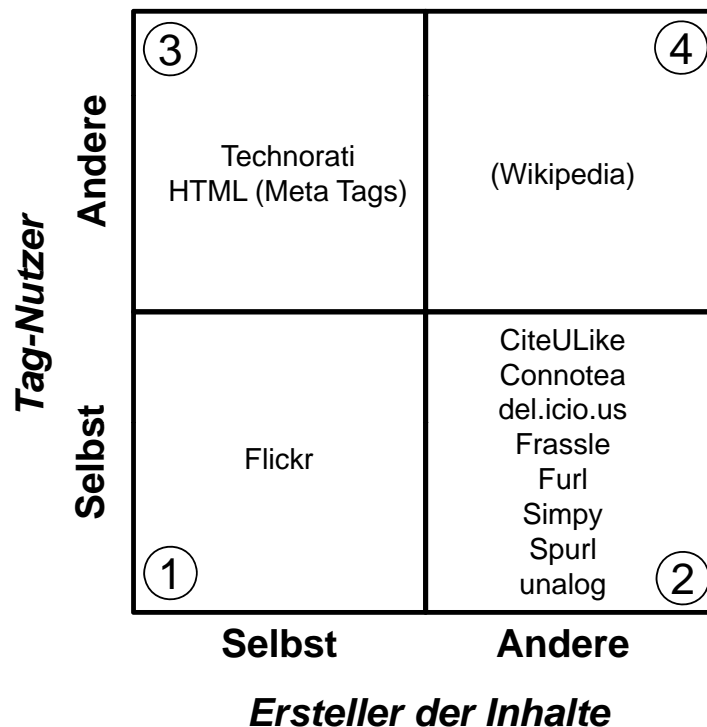


Abbildung 3.1: Akteure und Adressaten beim Tagging (vgl. [56])

Delicious ist hierfür ein Beispiel. Die Tags können auch für andere Adressaten interessant sein, aber das ist nicht das primäre Ziel für die meisten Nutzer.

Feld 3 - **Selbst-Andere:** Autoren, die ihre eigenen Inhalte für andere taggen, können zu dieser Gruppe zugeordnet werden. Ein Beispiel hierfür ist die Weblog-Suchmaschine Technorati⁶, mit denen Autoren ihren Weblog mit dem primären Ziel verschlagworten, dass andere Nutzer den Weblog finden. HTML-Meta-Tags stammen noch von vor der Zeit des heutigen Taggings, aber im Grunde genommen können Autoren mittels Meta-Tags Themen als Schlagworte in der Webseite hinzufügen, damit andere Nutzer die Webseite mittels Suchmaschine finden können.

Feld 4 - **Andere-Andere:** Diese Gruppe bezieht sich auf Nutzer, die fremde Inhalte für den Nutzen von anderen taggen. Als Beispiel wird an dieser Stelle Wikipedia genannt. In Wikipedia werden die Artikel zwar nicht mittels unkontrollierten Vokabulars getaggt, deshalb steht Wikipedia in Klammern, aber es werden den Artikeln Kategorien zugeordnet, die von der Community gepflegt werden. Diese Gruppe der Akteure, die fremde Inhalte verschlagwortet, wird in Abbildung 3.2 weiter unterteilt.

In Abbildung 3.2 werden die Akteure unterteilt in die Autoren, professionellen Indexer und Nutzer [75]. Seit Menschen Objekte organisieren, nutzen sie dafür Metadaten. Metadaten sind "Daten über Daten" [46]. In Bereichen wie Museen, Bibliotheken, Archiven usw. werden Metadaten genutzt, um Informationen zu organisieren und den Zugriff für andere zu ermöglichen. Akteure in diesen Bereichen sind in der Regel professionelle Indexer (in der Abbildung 3.2 als Gruppe 1 markiert). Professionelles Indexieren erfordert häufig den Gebrauch von standardisierten Wissensordnungen wie Nomenklaturen,

⁶ <http://technorati.com/>, Zuletzt abgerufen am 02.01.2011

Thesauri oder Klassifikationssystemen. Die verschiedenen *Wissensordnungen* werden in Abschnitt 3.2 definiert.

Eine häufig getroffene Anforderung für Metadaten ist, dass sie sowohl von Menschen lesbar als auch von Maschinen verarbeitbar sein sollen [66]. Im Bibliothekswesen gibt es hierfür Standards wie beispielsweise die Regeln für die alphabetische Katalogisierung (RAK) oder Dublin Core [69]. Im E-Learning existieren Metadatenstandards, wie Learning Objects Metadata (LOM) [66], zur Beschreibung von Lernressourcen. LOM enthält beispielsweise 9 Hauptkategorien mit insgesamt mehr als 70 Metadatenfeldern.

Das Ausfüllen von Formularen zur Metadatenerstellung ist häufig sehr aufwendig, komplex, fehleranfällig und setzt Expertenwissen bzw. Kompetenzen voraus, u.a. [12, 23, 26, 138, 62]. Diese Aussage trifft besonders auf E-Learning zu, da im E-Learning keine Experten des Bibliothekswesens, sondern häufig die Autoren der Inhalte die Metadaten zu ihren Inhalten erstellen. Die Autoren der Inhalte bilden die zweite Gruppe der Akteure.

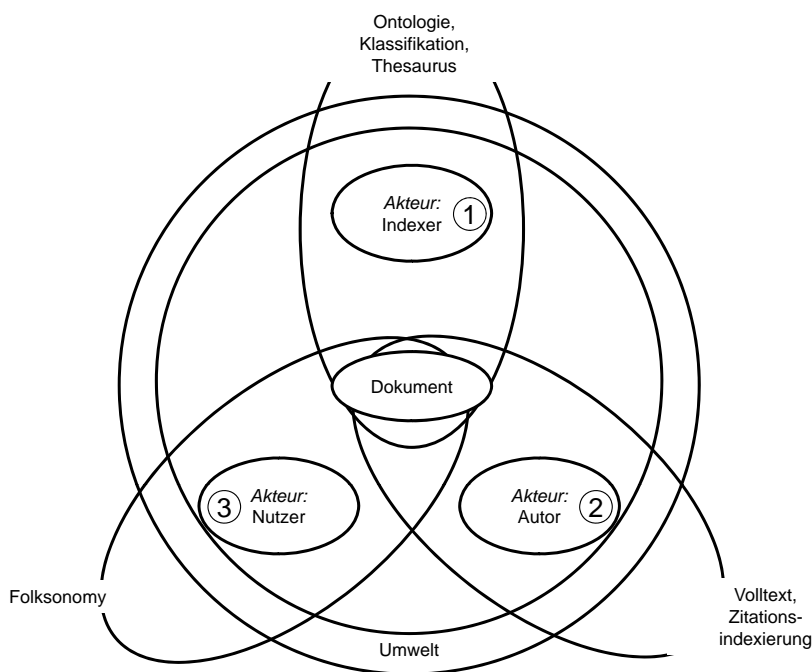


Abbildung 3.2: Methoden der Wissensrepräsentation und deren Akteure (vgl. [145])

Tagging, das von den Nutzern durchgeführt wird, führt zur dritten Gruppe und reduziert die Probleme der standardisierten Metadaten-Erstellung. Denn Tagging reduziert den Aufwand und die Komplexität der Erstellung von Metadaten und es wird kein Expertenwissen vorausgesetzt.

Da im Web die Menge an verschiedenen Informationen sehr groß ist, kann hier eine Metadaten-Erstellung durch wenige professionelle Indexer nicht durchgeführt werden. Der Aufwand des Indexierens kann durch maschinelle Verfahren reduziert werden. Allerdings ist die Qualität der maschinell erstellten Metadaten in der Regel geringer als bei manuellen Verfahren. Kommen rein textorientierte Verfahren zum Einsatz, die sich nur auf den Inhalt der Ressource beziehen, beschränken sich die Metadaten auf den Sprachgebrauch des Autors. Nutzer, die frei taggen, erweitern die Beschreibung der Informationen durch ihre eigene Sprache und zusätzliche Informationen wie Kontext-Informationen, die sich beispielsweise

auf die Verwendung der Informationen durch die Nutzer beziehen. Diese Informationen sind den Autoren und professionellen Indexern häufig unbekannt. Daher erstellen diese beiden Gruppen von Akteuren keine nutzerzentrierten Metadaten, d.h. keine Metadaten, die für den einzelnen Nutzer personalisiert sind. Die Gesamtheit der Tags in einem System wird meist als Folksonomie bezeichnet. Auf diese Form der Wissensorganisation geht der Abschnitt 3.2.6 ein.

Ein Ziel dieser Arbeit ist es, Akteure dabei zu unterstützen, fremde Inhalte im Web zu ihrem eigenen Nutzen zu taggen (siehe Abbildung 3.1 Feld 2 und Abbildung 3.2 Feld 3). Dienste, die das Tagging dieser Kategorie bereits unterstützen, werden als Social Bookmarking-Dienste bezeichnet und in Kapitel 6.4 analysiert. Der nächste Abschnitt gibt einen Überblick über die verschiedenen Funktionen von Tags.

3.1.3 Funktionen von Tags

Zahlreiche Studien untersuchen die Motivation der Nutzer beim Taggen und welche Funktionen Tags haben. Dieser Abschnitt gibt einen kurzen Einblick in die Motive und Funktionen. Die Motive hinter und die Funktionen von Tags führen zu verschiedenen Tag-Typen, die in Kapitel 4 aufgegriffen werden.

Das Ergebnis einer Studie von Marlow et al. [90] sind sechs Motive, die sich gegenseitig nicht ausschließen, d.h. einige dieser Motive können von den Nutzern gleichzeitig verfolgt werden.

- **Zukünftiges Wiederfinden:** Objekte werden mit Tags beschrieben, die beim späteren Finden und Erinnern helfen.
- **Beitragen und Austauschen:** Objekte werden getaggt, um die Sammlung der Ressourcen anderen Nutzern zur Verfügung zu stellen, z.B. eine Sammlung von Webseiten mit Ferienunterkünften.
- **Aufmerksamkeit erregen:** Andere Nutzer sollen auf die eigenen Objekte aufmerksam werden. Dazu werden Objekte mit häufig verwendeten Tags verschlagwortet.
- **Spiel und Wettbewerb:** Es gibt Nutzer bzw. Nutzergruppen, die Sichtbarkeit im System erzeugen, indem sie alle Objekte mit einem speziellen Tag versehen. Des Weiteren gibt es Tags, die von Nutzern nur eingefügt wurden, um Sätze innerhalb von Tag-Clouds (siehe Abschnitt 6.4) entstehen zu lassen.
- **Selbstdarstellung:** Die eigene Identität wird im System als "Markierung" hinterlassen, indem der eigene Bezug zum Objekt hergestellt wird, z.B. das Tag "live gesehen" in einem Video-System für ein Konzertvideo stellt einen individuellen Bezug zum Ersteller des Tags her.
- **Meinungen ausdrücken:** Tags werden genutzt, um Meinungen auszudrücken.

Golder und Huberman [47] identifizieren sieben verschiedene Funktionen von Tags. Dabei führen die ersten drei Funktionsbereiche zu objektiven Tags, d.h. zu Tags, die sich auf das getaggte Objekt beziehen und vom Nutzer relativ unabhängig sind. Die anderen vier Funktionsbereiche resultieren in persönliche Tags, die sehr subjektiv sein können und primär dem Nutzen des Tag-Erstellers dienen.

- **Thema:** Diese Tags identifizieren, von was oder wem das Objekt handelt.
- **Typ:** Diese Tags identifizieren den Typ des Objekts, z.B. "Artikel", "Blog", "Buch".
- **Autor oder Besitzer:** Mit diesen Tags wird der Ersteller oder der Besitzer des Objektes identifiziert.
- **Verfeinerungskategorien:** Dies sind individuelle Tags, die für sich allein genommen keinen Sinn ergeben, sondern die genutzt werden, um existierende Kategorien zu verfeinern (z.B. Tags wie

"25" und "100" erfüllen die Funktion, eine Kategorie weiter zu unterteilen um Unterkategorien zu bilden).

- **Bewertungen:** Meinungen des Nutzers werden in dieser Funktionsgruppe ausgedrückt, z.B. "lustig" oder "interessant".
- **Selbstreferenz:** Diese Tags dienen dazu, die Relation zwischen Objekt und Nutzer auszudrücken. Häufig beginnt das Tag mit "my", z.B. "mystuff" oder "mycomments".
- **Aufgabenmanagement:** Beim Sammeln von Objekten für eine bestimmte Aufgabe werden Tags zum Aufgabenmanagement genutzt, beispielsweise die Tags "toRead" oder "jobsearch".

Beim Tagging zeigt sich eine Vermischung von persönlichen Mehrwerten und implizitem Nutzen für die Community, obwohl der Nutzen für die Community beim Tagging häufig nicht im Vordergrund steht, sondern der persönliche Nutzen [81]. Anhand der Beispiele für Tagging-Systeme (siehe Abbildung 3.1) zeigt sich dieser Sachverhalt ebenso. In [58] wird die Motivation in zwei große Funktionsbereiche eingeteilt, zum einen in das persönliche Informationsmanagement und zum anderen in den Austausch von Ressourcen (Gruppenbezogenes Informationsmanagement). [102] erweitert die soziale Sicht auf die Wissensorganisation um die soziale Sicht auf die Kommunikationsfunktion von Tagging, siehe Abbildung 3.3, in der die resultierenden vier Sichten dargestellt sind.

		Funktionale Sicht	
		Organisation	Kommunikation
Soziale Sicht	Individuum	Individuelle Wissensorganisation: Inhalte suchen, kategorisieren und (wieder) auffinden	Individueller Interpretationshintergrund: Eigene Ideen, Erinnerungen und Kontakte dokumentieren
	Gruppe	Gruppenbezogene Wissensorganisation: Mit anderen Ressourcen teilen und/oder Aufmerksamkeit auf eigene Inhalte lenken	Sozialer Interpretationshintergrund: Beschreibungen und Wertungen zu Inhalten abgeben, Gruppenzusammenhalt ausdrücken (soziale Kohäsion)

Abbildung 3.3: Motive für die Nutzung von Tagging-Diensten (vgl. [102] basierend auf [6])

- **Individuelle Wissensorganisation:** Dieser Bereich umfasst das persönliche Wissensmanagement.
- **Gruppenbezogene Wissensorganisation:** Tags, die zum Austauschen von Objekten dienen, fallen in diesen Bereich. Dazu gehört auch die Motivation, Aufmerksamkeit für eigene Inhalte zu erreichen.

- **Individueller Interpretationshintergrund:** Dieser Bereich umfasst Tags, die dazu dienen, eigene Ideen und Situationen festzuhalten. Beispielsweise werden Fotos mit den Namen der fotografierten Personen getaggt, um sich an Details zu erinnern.
- **Sozialer Interpretationshintergrund:** Nutzer taggen, um Informationen zum Kontext des Objektes, gemeinsame Erfahrungen oder Wertungen auszutauschen. Damit wird Gruppenzusammenhalt ausgedrückt.

3.1.4 Tagging versus Klassifizierung

In einer Bibliothek kann ein Buch nur einen physikalischen Standort im Bücherregal haben. Daher muss sich der Bibliothekar entscheiden, an welchem Standort das Buch einsortiert wird. Das gleiche gilt für hierarchische Dateisysteme in Computern, wie sie zum Zeitpunkt dieser Arbeit noch weit verbreitet sind. In diesen können Dateien jeweils nur in einem Dateiordner gespeichert werden. Das Ablegen in verschiedenen Ordnern ist in diesen Fällen nur durch Kopien oder auch durch symbolische Verknüpfungen (sogenannten Softlinks) möglich, wodurch der Pflegeaufwand steigt. Beim Tagging jedoch sind die Benutzer in der Regel nicht in der Anzahl der Tags beschränkt, die pro Objekt vergeben werden dürfen. Die Benutzer müssen somit keine 1:1-Zuordnung vornehmen. Man spricht daher von *multiplen Ordnungssystemen*. Multiple Ordnungssysteme existieren nicht nur im Rahmen des Taggings, sondern beispielsweise auch in digitalen Bibliotheken. Allerdings können in digitalen Bibliotheken die Schlagworte in der Regel nicht so frei wie beim Tagging vergeben werden. Die Schlagworte können beispielsweise durch ein vorgegebenes Vokabular definiert sein und dann muss ein Abgleich mit dem Vokabular stattfinden oder die Anzahl der vergebenen Tags darf eine bestimmte Menge nicht überschreiten und dann muss dennoch eine Selektion der Schlagworte stattfinden.

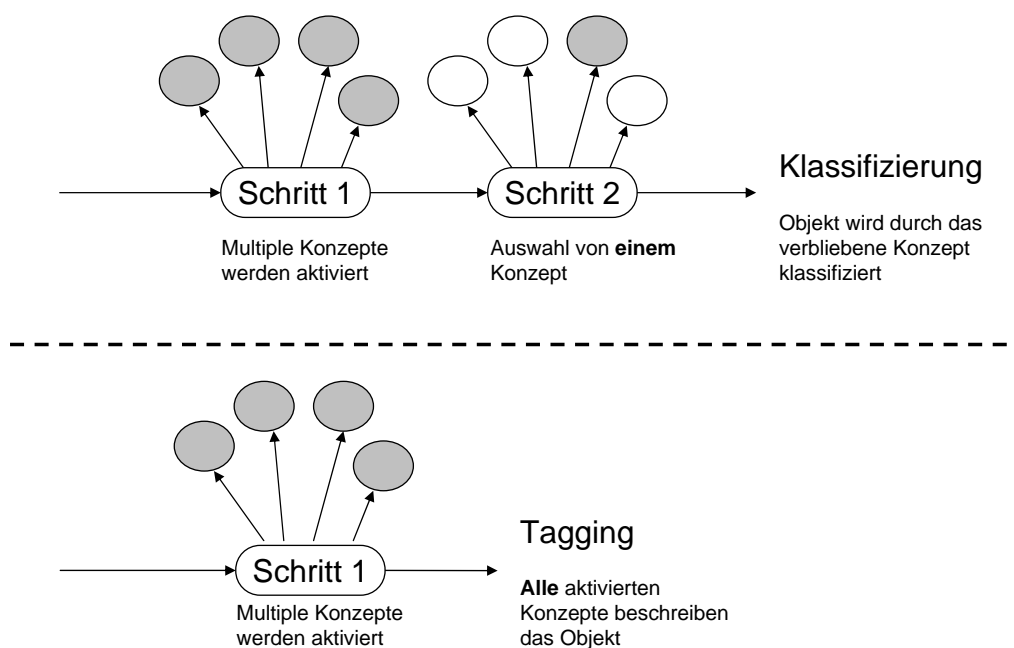


Abbildung 3.4: Kognitive Analyse des Tagging (vgl. [58] basierend auf [141])

Abbildung 3.4 veranschaulicht den Vorgang der Klassifizierung und des Taggings und deren kognitive Aufwände. Im ersten Schritt der Klassifizierung werden Schlagworte gesucht, die als Kandidaten in Frage kommen. Will ein Nutzer beispielsweise Webseiten, in denen das Thema Learning Object Metadata (LOM) behandelt wird, im Dateisystem speichern, könnten ihm die folgenden Ordnernamen in den Sinn kommen: "LOM", "E-Learning", "Metadaten", "Standard", "Referenz", "Projekt XY" etc. Im zweiten Schritt muss sich der Nutzer für einen dieser Begriffe entscheiden. Bei dieser Entscheidung ist wichtig, dass die Webseiten später wieder gefunden werden können. Dies verursacht einen kognitiven Mehraufwand verbunden mit der Unsicherheit, die richtige Kategorie ausgewählt zu haben. Beim Tagging fällt der zweite Schritt weg, d.h. die Entscheidung für eine einzige Klasse muss nicht getroffen werden. Tagging kann somit den kognitiven Aufwand reduzieren [141].

3.1.5 Zusammenfassung

Beim Tagging zeichnen Benutzer Objekte, wie beispielsweise Webseiten, mit unterschiedlichen Schlagworten aus, um die Objekte für das Finden bzw. Wiederfinden geeignet zu organisieren. Daher kann Tagging für das in dieser Arbeit beschriebene Ressourcenmanagement eingesetzt werden. Dieser Abschnitt des Kapitels gab eine Einführung in den Tagging-Vorgang, die Akteure und die möglichen Funktionen von Tags. Des Weiteren wurden Vorteile und Nachteile eines unkontrollierten Vokabulars herausgearbeitet. Diese Grundlagen werden in späteren Kapiteln wieder aufgegriffen.

3.2 Wissensrepräsentation

Verfahren zum Aufbau einer formalisierten Wissensbasis können im Gebiet der *Wissensrepräsentation* gefunden werden. Die Betrachtung der Wissensrepräsentation ist in dieser Arbeit eingeschränkt auf den Bereich der Verwaltung von elektronischen Dokumenten.

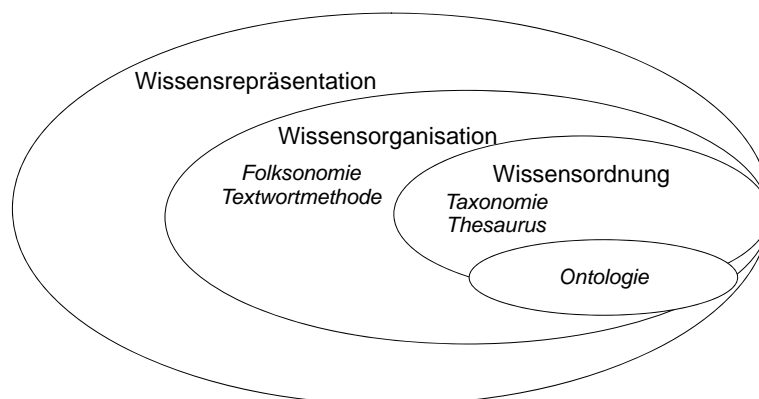


Abbildung 3.5: Begriffliche Einteilung Wissensrepräsentation, Wissensorganisation und Wissensordnung (vgl. [145])

Die Wissensrepräsentation ist ein Teilgebiet der Künstlichen Intelligenz. Nach [145] wird sie definiert als *"Wissenschaft, Technik und Anwendung von Methoden und Werkzeugen, Wissen derart abzubilden, damit dieses in digitalen Datenbanken optimal gesucht und gefunden werden kann. Sie ermöglicht die Gestaltung*

von Informationsarchitekturen, die - auf der Grundlage von Begriffen und Relationen arbeitend - gestalten, Wissen in seinen Bedeutungszusammenhängen darzustellen." Wissen wird demzufolge in eine für den Rechner verständlichen Form kodiert, damit das repräsentierte Wissen vom Rechner verarbeitet werden kann.

Die Verfahren der Wissensrepräsentation können weiter spezialisiert werden in Methoden der Wissensorganisation und diese wiederum in Methoden der Wissensordnung. Die Wissensordnung ist nach [145] der am engsten gefasste Begriff. Abbildung 3.5 erläutert den Zusammenhang dieser Begriffe.

Die Organisation von Wissen durch Begriffe wird als Wissensorganisation bezeichnet. Neben Methoden, die die Begriffe aus den Dokumenten extrahieren (z.B. Textwortmethode), gibt es auch die Möglichkeit des Taggings (Folksonomie). Wissensordnungssysteme sind Systeme, in denen sich zusätzlich die Begriffe ordnen lassen. Thesauri und Taxonomien sind Beispiele für Wissensordnungen. Die Wissensrepräsentation umfasst alle Wissensorganisationsmethoden.

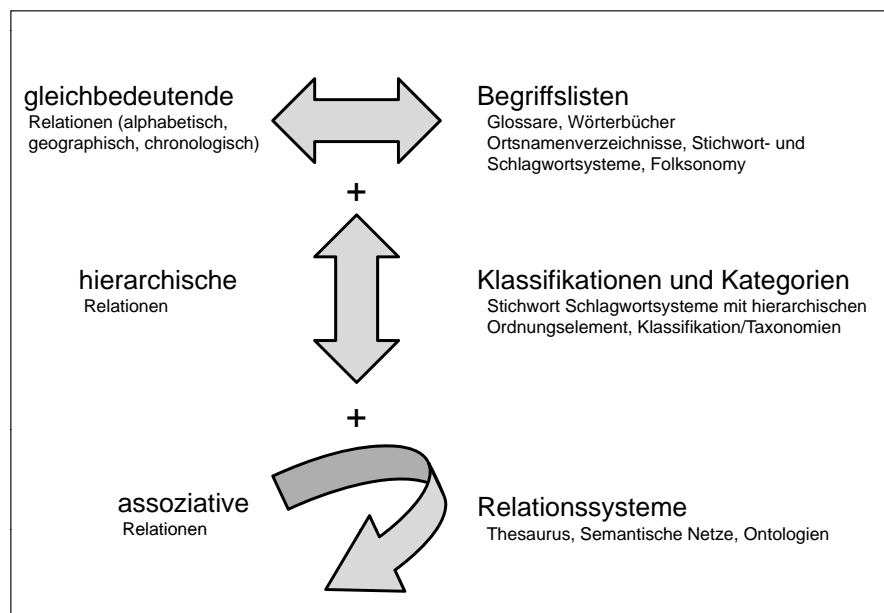


Abbildung 3.6: Relationstypen und Wissensorganisationssysteme (vgl. [45])

Eine feinere Einteilung der Wissensorganisationsmethoden kann entsprechend der möglichen Relationstypen vorgenommen werden [61], siehe Abbildung 3.6. Die Methoden gewinnen mit der Anzahl von unterschiedlichen Relationstypen an Ausdrucksstärke, aber auch an Komplexität.

- **Gleichbedeutende Relationen:** Begriffslisten beschreiben Ressourcen mit Begriffen, die mittels gleichbedeutender Relationen geordnet werden. Die Ordnung kann beispielsweise alphabetisch, geographisch oder chronologisch geschehen. Vertreter dieser Wissensordnung sind Glossare, Wörterbücher, Ortsnamenverzeichnisse, Stichwort- und Schlagwortsysteme oder Folksonomies (siehe Abschnitt 3.2.6).
- **Gleichbedeutende + Hierarchische Relationen:** Ordnungen, die zusätzlich hierarchische Relationen haben, gehören zu den Klassifikationen bzw. Kategorien. Sie teilen Begriffe in Klassen mit gemeinsamen Eigenschaften ein und ordnen diese Klassen hierarchisch. Hierunter fallen

Stichwort- und Schlagwortssysteme mit einem hierarchischen Ordnungselement und Taxonomien (siehe 3.2.1).

- **Gleichbedeutende + Hierarchische + Assoziative Relationen:** Relationssysteme ergänzen die Ordnung um assoziative Relationen. Thesauri (siehe 3.2.2), semantische Netze (siehe 3.2.4), Topic Maps (siehe 3.2.5) und Ontologien (siehe 3.2.3) sind Beispiele für Relationssysteme.

Im Folgenden werden verschiedene Verfahren der Wissensrepräsentation erläutert und anschließend hinsichtlich ihres Einsatzes für das Anwendungsszenario dieser Arbeit bewertet.

3.2.1 Taxonomien

Häufig wird der Begriff Taxonomie synonym zu Klassifikation verwendet. Eine Taxonomie gruppiert Dinge oder Begriffe anhand von gemeinsamen Merkmalen in Klassen. Eine Klasse kann dabei Oberklasse für mehrere Klassen sein. Ob ein Begriff mehrere Oberbegriffe haben darf, hängt davon ab, ob es sich um eine monohierarchische oder polyhierarchische Struktur handelt. Eine monohierarchische Struktur bildet eine Baumstruktur ab. Sie hat allerdings das Problem, dass eine Zuordnung in nur eine Klasse manchmal nicht eindeutig vorgenommen werden kann. In der Biologie sind Taxonomien weit verbreitet; sie dienen beispielsweise zur Erfassung der Artverwandtschaften von Lebewesen.

3.2.2 Thesauri

Neben der hierarchischen Ordnung von Begriffen wie in Taxonomien, werden in Thesauri die Begriffe durch weitere Äquivalenz- und Assoziationsrelationen vernetzt. Zu Äquivalenzrelationen zählen Synonyme, Übersetzungen, Abkürzungen und unterschiedliche Schreibweisen (z.B. Geographie/ Geografie). Assoziationsrelationen dienen dazu, den Bezug von verwandten Begriffen zueinander abzubilden. In der Dokumentationswissenschaft gibt es Normen wie die DIN 1463-1 [1] und die ISO 2788 [67], die festlegen, welche Relationstypen innerhalb von Thesauri erlaubt sind.

3.2.3 Ontologien

Der Begriff Ontologie kommt aus dem Griechischen von "on" für "das Seiende" und "logos" für "Lehre". Es bedeutet die *"Lehre vom Sein, von den Ordnungs-, Begriffs- und Wesensbestimmungen des Seienden"* [38]. Der Begriff wird in den Fachdisziplinen unterschiedlich definiert. Dieser Abschnitt beschränkt sich auf die Untersuchung von Definitionen in der Informatik.

Nach [143] wird eine Ontologie definiert als das Ergebnis *"of a study of the categories of things that exist or may exist in some domain"*. Im Fokus dieser Definition steht die gemeinschaftliche Studie der Kategorien (Unter-, Oberbegriffsrelation) von Dingen.

Eine andere weit verbreitete Definition von Ontologien stammt von Gruber [50]. Er definiert eine Ontologie als *"an explicit specification of a conceptualization"*, d.h. eine Ontologie ist eine explizite Spezifikation einer Begriffsbildung. Häufig wird *"Conceptualization"* aus dem Englischen in Konzeptualisierung übersetzt, damit werden die Termini "Begriff" und "Konzept" allerdings als Synonyme gebraucht. Panyr [103] und Faatz [42] kritisieren die Übersetzung in "Konzepte" statt "Begriffe" für Ontologien.

Konzepte sind Vorstellungen, die Wörter im menschlichen Denken hervorrufen. Begriffe sind Abstrahierungen dieser Konzepte. In [54] werden folgende Unterscheidungen getroffen: Ein Konzept ist *"jede Regel, nach der bestimmte Reize mit einer Reaktion verknüpft werden"*. Ein Begriff ist *"eine durch das Denken gewonnene, umgrenzte Allgemeinvorstellung, in der eine Summe von Einzelvorstellungen zusammengefasst ist"*. Begriffe können demnach als Menge ihrer Eigenschaften definiert werden (Begriffsintension/ Begriffsinhalt) oder als Menge ihrer zugehörigen Elemente (Begriffsextension/ Begriffsumfang) [42]. Jeder Begriff muss eindeutig bezeichnet werden. Daraus folgt, dass jeder Begriff eindeutig zu einem Konzept führen muss. Beispielsweise muss der Begriff der "Ente", der das Tier bezeichnen soll, eindeutig zum Konzept des Tieres führen und darf nicht das Konzept der Zeitungsentente ableiten. "Jedoch ist die Existenz eines Konzepts nicht an die sprachliche Bezeichnung gebunden, und es sind im Handeln viele Konzepte wirksam, denen ein sprachlicher Ausdruck fehlt" [54]. Demnach kann nicht jedes Konzept durch einen eindeutigen sprachlichen Bezeichner ausgedrückt werden, z.B. "Wahrheit", "Freiheit" etc.

Gruber nennt in [50] zusätzliche Anforderungen an eine Ontologie. Sie soll formal sein, damit sie maschinenlesbar ist und sie soll einvernehmlich sein, da die Erfassung des geteilten Wissens einer Community möglich sein soll. In [147] werden diese Anforderungen in die Definition aufgenommen: *"An ontology is a formal, explicit specification of a shared conceptualisation."* Eine Ontologie ist demnach eine formale und explizite Spezifikation einer gemeinsamen Begriffsbildung.

In Ontologien werden in der Regel Begriffe und Instanzen unterschieden. Beispielsweise ist Anne eine Instanz des Begriffs Person. Die Trennung von Begriffen und Instanzen ist auch auf die Relationen übertragbar, sodass die möglichen Relationen von den tatsächlich existierenden Relationen unterschieden werden können. Beispielsweise können zwischen zwei Personen die Relationen "ist verheiratet mit" und "ist verschwägert mit" mögliche Relationen darstellen, wobei zwischen den Instanzen Anne und Dirk nur die Relation "ist verheiratet mit" tatsächlich existiert.

Ontologien verfügen nicht nur über eine Begriffsordnung, sondern auch über *Axiome* und Mechanismen zum *automatischen Schließen*. Daher haben Ontologien eine Sonderstellung gegenüber den anderen Wissensorganisationsmethoden wie Taxonomien und Thesauri. Automatisches Schließen erfolgt durch logische Regeln, indem neues Wissen von existierendem Wissen abgeleitet wird, z.B. "Wenn das und das gilt, dann gilt auch jenes". Aus "Vögel können fliegen" und "Flori istEin Vogel" folgt "Flori kann fliegen". Axiome sind Aussagen die immer wahr sind und nicht aus anderen Bestandteilen der Ontologie hervorgehen bzw. geschlussfolgert werden können.

Die Erstellung von Ontologien ist in der Regel ein formaler Prozess mit hohem Kommunikationsaufwand, der beispielsweise durch intensive Diskussionen zur Begriffsbildung verursacht wird [42]. Ontologien sind eine sehr mächtige, aber auch sehr komplexe Methode, um Wissen zu repräsentieren.

3.2.4 Semantische Netze

Ein semantisches Netz ist nach [142] *"a graphic notation for representing knowledge in patterns of interconnected nodes and arcs"*, d.h. eine graphische Darstellung von Wissen in Form von verbundenen Knoten und Kanten. Der Ursprung von semantischen Netzen geht auf Quillian [110] zurück. Sein Ziel war es, eine reichhaltigere Repräsentation von Wortbedeutungen in Thesauri zu ermöglichen.

Im Allgemeinen werden die Knoten und Kanten beschriftet. Die Knoten werden als Konzepte interpretiert. Bei semantischen Netzen darf im Gegensatz zu Ontologien auch von Konzepten gesprochen werden [42] (siehe Unterscheidung Begriff und Konzept in Abschnitt 3.2.3). Die Kanten werden in semantischen Netzen auch als Relationen bezeichnet. Die kleinste Informationseinheit sind zwei durch eine Relation verbundene Konzepte. Jedes Konzept wird durch die Verbindung zu anderen Konzepten definiert. Unverbundene Konzepte haben keine Bedeutung.

Quillian unterscheidet die Knoten in Token und Typen, z.B. Anne ist ein Token und Person ist ein Typ. Dies kann analog zur Unterscheidung von Instanz und Begriff in Ontologien betrachtet werden. Allerdings wird in semantischen Netzen diese Unterscheidung durch eine benannte Relation dargestellt z.B. Anne "ist vom Typ" Person. Ebenso können die Typen durch spezielle "Subtyp"-Relationen geordnet werden z.B. Person "ist Subtyp von" Lebewesen. Die Existenz und die Notwendigkeit solcher "Subtyp"-Relationen als Basisstruktur zur Ordnung von Konzepten ist in der Literatur zu semantischen Netzen im Gegensatz zu Ontologien nicht deutlich definiert [42, 65]. Mitunter wird die Ordnung der Typen als Schema ausgelagert. Das Schema umfasst alle möglichen Typen und deren mögliche Relationen.

Die Ordnung von Begriffen und Relationen übernimmt Beier in seine Definition [14] und bezeichnet diese als Typisierung: *"Ein semantisches Netz ist eine geordnete Zusammenstellung von Begriffen und ihren Bezeichnungen, deren Zusammenhang über beliebige Beziehungen miteinander definiert wird. Sowohl Begriffe als auch Beziehungen sind typisierbar und es existiert eine Grammatik für deren Verwendung."* Beier verwendet in seiner Definition den Begriffs-Begriff statt des Konzept-Begriffs, obwohl eine formale Begriffsbildung in semantischen Netzen nicht gefordert wird.

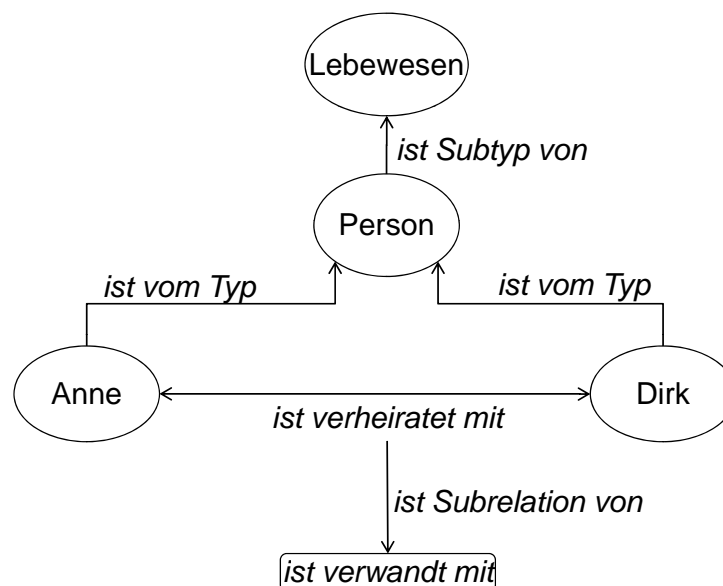


Abbildung 3.7: Beispiel für ein semantisches Netz mit Typisierung

Definition 3 Semantische Netze werden somit definiert als eine Ansammlung von Knoten, die durch beliebige Relationen in Beziehung miteinander stehen. Die Knoten können Token oder Typen darstellen. Sowohl Token, Typen und Relationen können typisiert werden.

Ein Beispiel für ein semantisches Netz mit Typisierung der Token, Typen und Relationen ist in Abbildung 3.7 dargestellt. Anne und Dirk sind Token vom Typ Person und Person ist Subtyp von Lebewesen. Die Typisierung von Typen führt zur Bildung von hierarchischen Beziehungen zwischen den Typen. Anne und Dirk sind durch die Relation "ist verheiratet mit" verbunden, die Relation wiederum ist Subrelation von "ist verwandt mit".

In semantischen Netzen können, ähnlich zu Ontologien, Ableitungsregeln für das logische Schließen modelliert werden.

3.2.5 Topic Maps

Topic Maps sind durch den ISO-Standard ISO/IEC 13250 [68] definiert. Sie bestehen, wie die anderen Relationssysteme auch, aus Konzepten, die durch Relationen verbunden werden. Die Konzepte werden *Topics* genannt und repräsentieren Themen oder Dinge. Ein Topic darf mehrere Benennungen haben, die durch sogenannte *Scopes* einen bestimmten Gültigkeitsbereich besitzen. Beispielsweise können die Benennungen in verschiedenen Sprachen vorliegen und der Scope bestimmt, wann welche Benennung gilt. Es kann demzufolge Topics mit demselben Namen geben, aber in einem Scope muss jeder Name eindeutig sein. Beispielsweise kann es "Java" einmal als Insel im Scope "Geographie" und einmal als Programmiersprache im Scope "Informatik" geben. Die Relationen werden als *Assoziationen* bezeichnet. Die Topics können in einer Assoziation bestimmte Rollen einnehmen, z.B. kann eine Person in der Assoziation zu ihrem Ehemann die Rolle der Ehefrau annehmen und in der Assoziation zu ihren Kindern die Rolle der Mutter. Dieses Assoziationsmodell ist sehr flexibel, aber auch sehr komplex. Die Topics, Assoziationen und Rollen können typisiert werden, wobei die Typen ebenfalls als Topics definiert werden, d.h. es gibt kein Spezial-Topic für Typen. Weitere Bestandteile in Topic Maps sind Verlinkungen von Topics mit Informationsressourcen wie Webseiten, sie werden als *Occurrences* bezeichnet.

Topic Maps können als semantische Netze betrachtet werden, die zusätzlich Dokumente als *Occurrences* in ihre Wissensbasis aufnehmen [42]. Denn die Topics können durch das Konstrukt der *Occurrences* getrennt von den Informationsressourcen organisiert werden, somit funktionieren die Topics auch eigenständig als semantisches Netz.

3.2.6 Folksonomien

Durch die Aggregation der Tags aller Nutzer in einem System entsteht eine Struktur, die weit verbreitet als Folksonomie (engl. Folksonomy) bezeichnet wird. Diesen Begriff hat Vander Wal im Jahr 2004 eingeführt, seine Erläuterungen zur Entstehung des Begriffes finden sich in [152]. Der Begriff ist zusammengesetzt aus "folk" und "taxonomy" und bezieht sich darauf, dass gemeinschaftliches Klassifizieren stattfindet. Streng genommen ist eine Folksonomie keine Taxonomie, da es in Tagging-Systemen keine hierarchischen Relationen gibt, sondern nur gleichberechtigte Relationen möglich sind. Ein weiterer Unterschied zur Taxonomie ist, dass Taxonomien in der Regel "top-down" erstellt werden und Folksonomien "bottom-up" [4]. Der Top-Down-Ansatz gibt eine Struktur vor, mittels derer die Ressourcen klassifiziert werden; wohingegen die Struktur bei Bottom-Up-Ansätzen erst durch die Klassifikation der Ressourcen entsteht.

Durch die Einfachheit der Relationen sinkt zwar die Komplexität des Tagging-Vorgangs, aber die Ordnung der Begriffe verliert an Mächtigkeit. Dies kann beispielsweise in späteren Suchen von Nachteil sein. Dieser Nachteil könnte durch den Einsatz von Ontologien gemindert werden. Allerdings hat eine Ontologie, die durch Experten erstellt wird, die bereits in Abschnitt 3.1.1 erwähnten Schwächen eines kontrollierten Vokabulars. Die Trennung von der Ontologie-Erstellung durch Experten und der Ontologie-Nutzung durch die Anwender führt häufig dazu, dass die Anwender unzufrieden mit der Ontologie sind [18]. Denn Experten haben zwar Expertise in der Erstellung von Ontologien, aber es fehlt ihnen häufig das Fachwissen der Domäne, die sie modellieren. Des Weiteren können Ontologien auf Grund eines langwierigen Erstellungsprozesses oder eines aufwendigen Pflegeprozesses nur schwer aktualisiert werden und dadurch schnell veraltet sein. Dadurch können Ontologien entstehen, die in den Augen der Nutzer fehlerhaft, unvollständig oder nicht nachvollziehbar sind.

Studien zeigen, dass eine Folksonomie keine chaotische Ansammlung von Tags ist, sondern dass sich in der Community häufig ein gemeinsames Verständnis und Vokabular entwickelt [47, 90, 134, 104]. Daher gibt es Ansätze, die versuchen, Ontologien aus Folksonomien zu entwickeln. Dies geschieht zum einen durch Verfahren, die die Anwender in den Erstellungsprozess integrieren, und zum anderen durch maschinelle Verfahren.

In [18] wird das Konzept eines Werkzeugs vorgestellt, mit dem die Nutzer am sogenannten Ontologiereifungsprozess durch Tagging aktiv teilnehmen können. Sie können neue Begriffe in Form von Tags vorschlagen. Die Verwendung dieser Tags in der Community führt zum gemeinsamen Vokabular. Weitere Strukturierungen durch hierarchische und assoziative Relationen können anschließend von der Community vorgenommen werden. Die Formulierung von Axiomen bleibt den Experten überlassen. In [53] wird ein Ansatz vorgestellt, bei dem ein Vokabular-Moderator neue Tags in das kontrollierte Vokabular aufnimmt und klassifiziert.

Neben manuellen Ansätzen gibt es Verfahren, die versuchen aus einer Folksonomie automatisch eine Ontologie (bzw. Taxonomie) zu generieren, z.B. [63, 165, 85, 28, 94, 31]. Auf diese Verfahren wird an dieser Stelle nicht näher eingegangen, da die Generierung von Ontologien nicht Ziel dieser Arbeit ist.

3.3 Fazit

Für die Unterstützung des Ressourcen-basierten Lernens wird ein System zur Organisation der Ressourcen, der Informationen und des Wissens benötigt. Dieses Kapitel betrachtete grundlegend den Vorgang des Taggings zur Organisation von Webressourcen. Weitere Formen der Wissensrepräsentation wurden vorgestellt.

Die Begriffe Ontologie, semantisches Netz und Topic Map werden in der Informatik häufig synonym gebraucht. Die verschiedenen Formen der Wissensrepräsentation können ineinander überführt werden. Dies kann auf Grund der unterschiedlichen Mächtigkeit der Relationssysteme jedoch nicht immer verlustfrei erfolgen, somit verliert die Wissensrepräsentation ggf. an Bedeutung. Die Möglichkeit zur Definition von Axiomen fehlt beispielsweise in semantischen Netzen und Topic Maps. Alle Strukturen lassen sich graphisch als Netz in Form von Knoten und Kanten darstellen. Aus Sicht des Anwenders sind die Unterschiede der Formen der Wissensrepräsentation zumeist unerheblich. Oft werden Ontologien als Spezialfälle von semantischen Netzen bezeichnet [42, 144]. Diese Arbeit teilt dieses Verständnis und betrachtet semantische Netze als Oberbegriff.

Im Anwendungsszenario dieser Arbeit modellieren Anwender ihr Wissen. Diese Anwender sind keine Experten in der Erstellung von Wissensrepräsentationen. Der Erstellungsprozess von formalen Wissensrepräsentationen mit der höchsten Ausdrucksstärke wäre daher zu aufwendig. Des Weiteren sind Anwender meist nicht in der Lage, Axiome zu definieren. Da keine explizite Begriffsbildung stattfindet, kann streng genommen nicht von Ontologien gesprochen werden.

Tagging-Systeme erlauben den Aufbau einer Wissensstruktur durch Verschlagwortung von Objekten wie Webseiten. Allerdings ist die Ausdrucksmöglichkeit beim Tagging durch die Einfachheit stark eingeschränkt. Daher wird als Lösung ein Tagging-System vorgeschlagen, das durch die Möglichkeit, hierarchische und assoziative Relationen zu erfassen, an Mächtigkeit gewinnt, ohne dabei die Leichtigkeit des Taggings (zu stark) aufzugeben.

Für die folgenden Kapitel wird der Begriff "Wissensnetz" eingeführt, um von den Begriffen "Ontologie" und "semantisches Netz" abzugrenzen und Missverständnissen vorzubeugen. Ein Wissensnetz bezeichnet hierbei die Wissensstruktur der Anwender, die durch das Tagging und die Wissensmodellierung entsteht. Des Weiteren findet im Anwendungsszenario des Ressourcen-basierten Lernens keine formale Begriffsbildung statt, so dass im Folgenden immer der Konzept-Begriff verwendet wird.

Folgende Anforderungen werden im Anwendungsszenario an die Wissensrepräsentation gestellt:

- Die Wissensmodellierung darf kein Expertenwissen erfordern.
- Die Wissensmodellierung muss individuell und flexibel sein.
- Die Lernenden müssen sich leicht im Wissensnetz zurechtfinden können.
- Es muss eine geeignete Visualisierung des Wissensnetzes geben.

4 Typisiertes Tagging im Ressourcen-basierten Lernen

Eine stärkere Ausdrucksmöglichkeit in Tagging-Systemen wird durch semantisches Tagging ermöglicht. Eine Form von semantischem Tagging ist typisiertes Tagging, bei dem der Typ des Tags angegeben wird. In diesem Kapitel wird das Konzept des semantischen Taggings erläutert und es werden ausgewählte semantische Tagging-Systeme beschrieben. Anschließend wird eine Basismenge aus Tag-Typen für das Ressourcen-basierte Lernen anhand der Literatur erarbeitet und anhand der Ergebnisse von durchgeführten Studien zur Prüfung der Vollständigkeit dieser Basismenge diskutiert. Des Weiteren wird ein Verfahren zur automatischen Erkennung dieser Tag-Typen vorgestellt und evaluiert. Typisierte Tags können zur Verbesserung von Empfehlungen von Ressourcen und Tags in Tagging-Systemen dienen. Dies wird im letzten Abschnitt dieses Kapitels erläutert.

4.1 Semantisches Tagging

In diesem Abschnitt wird semantisches Tagging näher erläutert und anschließend durch konkrete Tagging-Systeme demonstriert.

4.1.1 Definition

Tags haben im Allgemeinen eine bestimmte, oft implizite, semantische Bedeutung und Funktion für den Nutzer, der die Ressource taggt. Ein Überblick über die verschiedenen Funktionen von Tags findet sich in Kapitel 3.1.3. So werden beispielsweise Tags verwendet, die den Inhalt der Ressource beschreiben, oder die die eigene Meinung über die Ressource ausdrücken wie "interessant" oder "lustig". Manche Tags beziehen sich auf die Aufgabe, in deren Rahmen die Ressource gefunden wurde. Solche Tags sind nur für den taggenden Nutzer sinnvoll, da sie sich auf eine konkrete, persönliche Situation beziehen. Menschen können anhand der Ressource und weiteren Tags, mit denen die Ressource getaggt wurde, in vielen Fällen die implizite Bedeutung des Tags verstehen. Ist eine Ressource beispielsweise mit "Java" getaggt, die ein Tutorial zum Thema "Programmieren in Java" enthält, kann ein Mensch verstehen, dass mit dem Tag "Java" die Programmiersprache und nicht die Insel oder der Kaffee gemeint ist.

Ein Computer kennt die Bedeutung der Tags nicht, für ihn sind Tags erst einmal nur Zeichenketten. Wird in einem Tagging-System nach dem Tag "Java" gefiltert, dann werden alle Ressourcen, die mit diesem Schlagwort getaggt wurden, zurückgeliefert - gleich ob der Nutzer die Programmiersprache oder die Insel meinte. Eine Möglichkeit, die Bedeutung des Tags auszudrücken, ist die Verwendung eines kontrollierten Vokabulars, in dem jeweils ein Begriff für Programmiersprache, Insel und Kaffee definiert wird. Dies wiederum ist mit den in Kapitel 3.1.1 genannten Nachteilen eines kontrollierten Vokabulars verbunden.

Semantisches Tagging¹ versucht die Lücke zwischen freiem und kontrolliertem Vokabular zu schließen, indem es nicht nur den Tag-Bezeichner, sondern auch dessen Semantik erfasst, d.h. die Bedeutung

¹ Semantisches Tagging bezieht sich in der Linguistik vor allem auf die Auszeichnung von Objekten, Subjekten etc. in natürlichsprachigen Texten. Dieser Abschnitt beschäftigt sich aber mit der semantischen Auszeichnung von Tags, mit denen Ressourcen verschlagwortet werden.

eines Tags wird, wie in [105] beschrieben, beispielsweise in Form einer URI, die auf ein Konzept aus einem definierten Vokabular verweist, zum Schlagwort hinzugefügt. Das Vokabular kann in einer externen Quelle definiert sein, z.B. wird die Wikipedia häufig als Vokabular genutzt; deren Artikel entsprechen dann den Konzepten und die URL der Artikel wird als URI verwendet. Beispielsweise kann die URI, die dem Tag "Java" als Bedeutung hinzugefügt wird, auf einen Artikel in der Wikipedia verweisen, der von der Programmiersprache Java handelt. Ein semantisches Tag entspricht dann einem n-Tupel der Form:

(Objekt, Tag-Bezeichner, Bedeutung, Benutzer, Zeitstempel, ...)

Das Tag-Tupel im konkreten Beispiel verknüpft mit einer Ressource über die Programmiersprache "Java" könnte dann die folgende Form haben:

*(<http://openbook.galileocomputing.de/javainsel/>, Java,
[http://de.wikipedia.org/wiki/Java_\(Programmiersprache\)](http://de.wikipedia.org/wiki/Java_(Programmiersprache)), Benutzer, Zeitstempel, ...)*

Als Basis für semantisches Tagging können Datenformate wie das Simple Knowledge Organisation System (SKOS) [163] oder das Tagging-Format Common Tag² dienen. SKOS ist ein Datenmodell zur Kodierung kontrollierter Vokabulare wie Thesauri. Das World Wide Web Consortium (W3C) stellt mit SKOS eine Empfehlung für ein Framework zur Veröffentlichung, zum Austausch und zur Integration von Daten aus verschiedenen Quellen bereit. Abbildung 4.1 erläutert ein Beispiel für die Verwendung von SKOS zur Formalisierung von Tags beim semantischen Tagging. In der Abbildung sind zwei getaggte Objekte (#FotoSparkasse und #FotoParkbank) zu sehen, die mittels der Relation *skos:subject* mit jeweils einem Konzept *skos:concept* (my:b und my:c) verknüpft sind. Mit der Relation *skos:prefLabel* und *skos:altLabel* können bevorzugte bzw. alternative Bezeichner festgelegt werden, wobei bevorzugte Bezeichner pro Vokabular eindeutig sein müssen. In SKOS wird somit eine Trennung von Konzepten und deren Bezeichnern realisiert. Im Beispiel ist "Sparkasse" der bevorzugte Bezeichner für das Konzept "my:b", das das #FotoSparkasse indexiert. "Sitzbank" ist für das Konzept an #FotoParkbank die bevorzugte Bezeichnung und "Bank" seine alternative Bezeichnung. Konzepte können mittels der Relationen *skos:broader* und *skos:narrower* hierarchisch untereinander verknüpft werden. Im Beispiel ist "my:b" ein Unterkonzept von "my:a". Durch diese semantische Annotation von Tags und durch die Verknüpfung von Tags untereinander kann eine Tagging-Struktur aufgebaut werden, die Wissensorganisationen wie beispielsweise Thesauri ähnlich wird.

Common Tag ist ein Tagging-Format, das ähnlich wie SKOS darauf zielt, Inhalte durch semantische Tags besser zu verknüpfen als dies mit freien Tags auf Basis von Zeichenketten möglich ist. Es werden hierbei nicht nur Nutzer ins Auge gefasst, die Webseiten in einem Tagging-System verschlagworten, sondern es sollen auch Autoren von Webseiten dabei unterstützt werden, ihre Inhalte durch semantische Tags zu beschreiben und untereinander zu verknüpfen. An der Entwicklung dieses Formats sind Unternehmen wie Yahoo³ und Google⁴ beteiligt. Semantische Tags sind in Common Tag Referenzen auf eindeutige Konzepte, die beispielsweise durch Einträge in vorhandenen, ein Vokabular bildenden Da-

² <http://commontag.org/Specification>, Zuletzt abgerufen am 09.02.2011

³ <http://www.yahoo.com/>, Zuletzt abgerufen am 09.02.2011

⁴ <http://www.google.com>, Zuletzt abgerufen am 09.02.2011

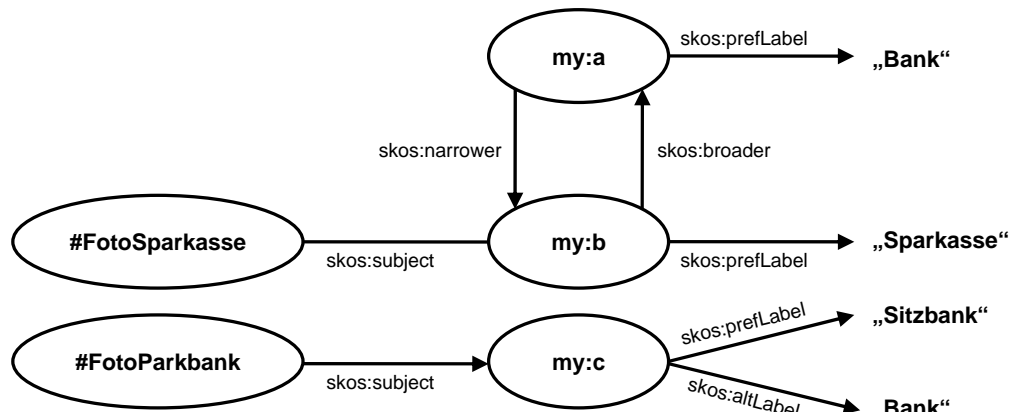


Abbildung 4.1: Beispiel zur Kodierung von semantischen Tags in SKOS vgl. [153]

tenquellen wie Freebase oder DBpedia⁵ definiert werden. Diese Einträge sind untereinander verlinkt und enthalten strukturierte Informationen. Beispielsweise enthält das Konzept über Barack Obama die Information, dass er Präsident der USA und mit Michelle Obama verheiratet ist. Freebase und DBpedia werden in Abschnitt 5.2.2 näher erläutert. Das Tagging-Format enthält weniger Elemente als SKOS. Abbildung 4.2 veranschaulicht Teile des Formats. Inhalt (*<resource>*), der durch eine URL adressiert werden kann, wird mittels Relation (*tagged*) mit einem Tag verknüpft, der mittels Relation (*means*) auf ein Konzept verlinkt, das wiederum durch eine URL (*<resource>*) beschrieben wird, die in der Regel auf einen Eintrag in der DBpedia verweist. Optional kann das Tag einen Bezeichner (*label*) und Zeitpunkt des Taggings (*taggingDate*) haben. Weitere Elemente sind Unterklassen von Tag: *AuthorTag* identifiziert ein Tag, das vom Autor der Ressource erstellt wurde, *ReaderTag* identifiziert ein Tag, das vom Konsumenten einer Ressource vergeben wurde und *AutoTag* spezifiziert ein Tag, das durch einen automatischen Tagging-Vorgang erstellt wurde.

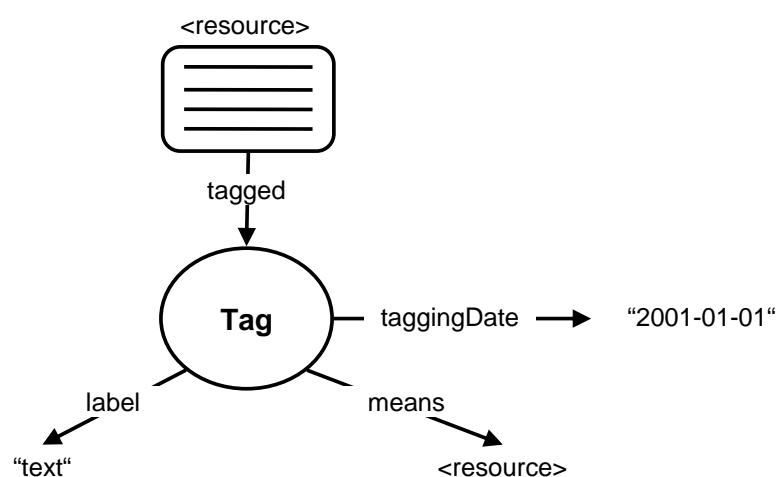


Abbildung 4.2: Common Tag-Modell⁶

⁵ <http://dbpedia.org/>, Zuletzt abgerufen am 09.02.2011

⁶ <http://commontag.org/QuickStartGuide>, Zuletzt abgerufen am 09.02.2011

4.1.2 Semantische Tagging-Systeme

Im Folgenden werden exemplarisch drei semantische Tagging-Anwendungen hinsichtlich der Erfassung von Semantik beim Tagging verglichen. Dazu werden die Aspekte der Semantik von Tags, der zusätzlich erfassbaren Informationen über Ressourcen und Tags, der möglichen Relationen zwischen Tags (z.B. Synonymrelationen) und deren Gültigkeitsbereich (beschränkt auf einen Benutzer oder benutzerübergreifend gültig) näher betrachtet. Weitere semantische Tagging-Systeme sind u.a. Bibsonomy [64], SemKey [89] und TagMe! [2].

Zigtag

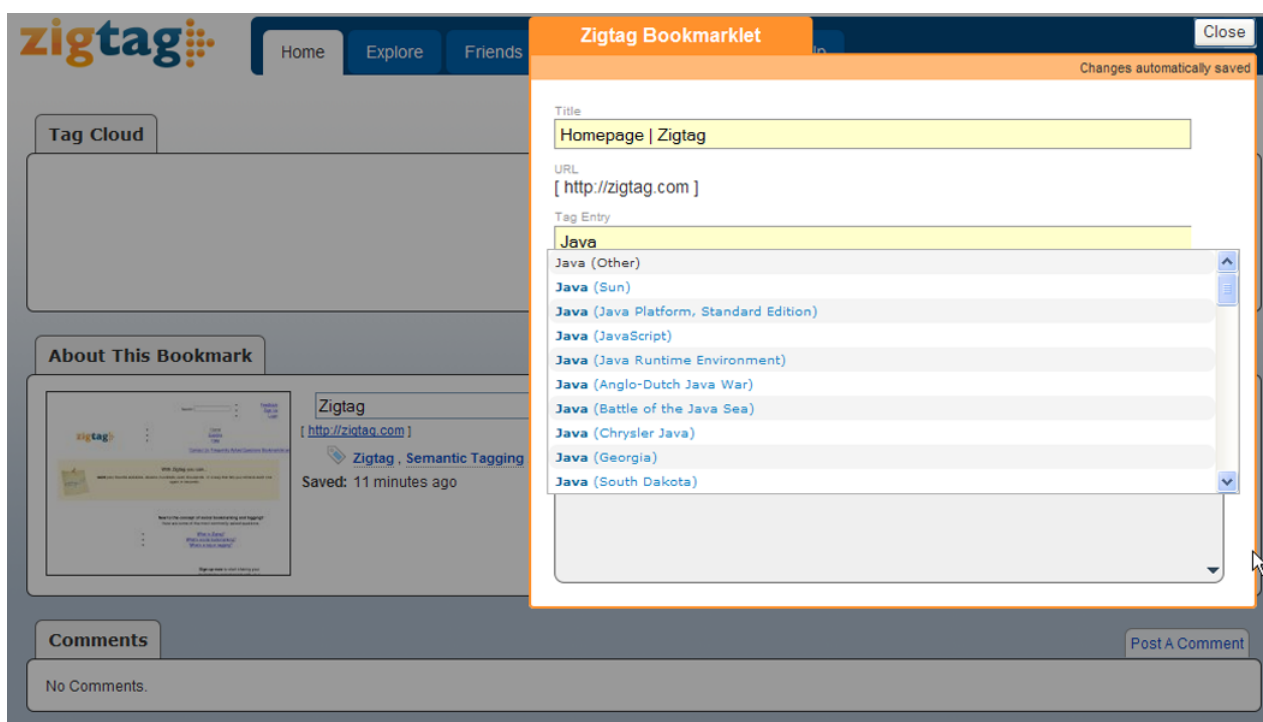


Abbildung 4.3: Tagging Dialog in ZigTag

Zigtag⁷ speichert Tags im Common-Tag-Format (siehe Abschnitt 4.1.1). Bei Eingabe eines Schlagwortes werden Konzepte aus den vorhandenen, ein Vokabular bildenden Datenquellen vorgeschlagen. In Abbildung 4.3 werden zur Eingabe von "Java" verschiedene Konzepte aufgelistet. Es können auch Tags verwendet werden, die keine Konzepte in der Datenquelle referenzieren. Diese werden als undefinierte Tags gekennzeichnet und können auch nach dem Tagging-Prozess vom Nutzer noch mit existierenden Konzepten in der Datenbank verknüpft werden. Definierte Tags tragen den Namen des Konzeptes als Bezeichner und können durch den Nutzer nicht umbenannt werden, wodurch Probleme durch Gleichbenennung auftreten können, denn der Bezeichner für Java als Insel und Java als Programmiersprache ist beispielsweise identisch. Der Unterschied kann in Zigtag beim Browsen nur durch die Einblendung von kurzen Beschreibungen der Konzepte aus der Datenquelle zu dem jeweiligen Tag sichtbar gemacht

⁷ <http://zigtag.com>, Zuletzt abgerufen am 09.02.2011

werden. Neben Titel und URL der Ressource können als Zusatzinformation Kommentare von Nutzern eingegeben werden. Relationen zwischen Tags sind nicht möglich. Ressourcen und deren Tags werden im Gültigkeitsbereich des jeweiligen Nutzers verwaltet. Die Definition der Tags aber ist auf Grund der verpflichteten Nutzung der Datenquelle global.

Faviki

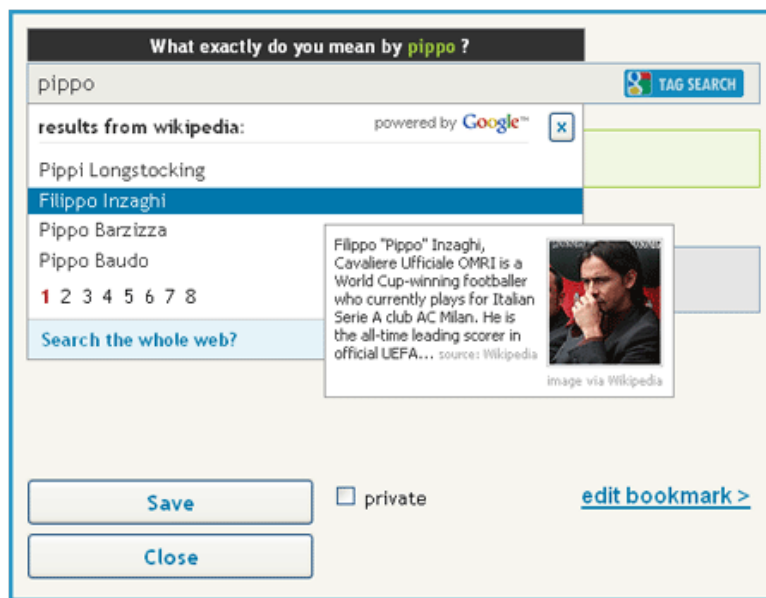


Abbildung 4.4: Tagging Dialog in Faviki

Ebenso wie Zigttag nutzt Faviki⁸ das Common-Tag-Format. Als Datenquelle zur Beschreibung der Tags werden Wikipedia und Google genutzt. Während der Eingabe eines Tags werden den Benutzern mittels DBpedia Artikel aus der Wikipedia vorgeschlagen, d.h. wird beispielsweise "pippo" eingegeben, kann der Benutzer u.a. den Wikipedia-Artikel über Pippi Longstocking oder den Wikipedia-Artikel über den Fußballspieler Filippo Inzaghi auswählen (siehe Abbildung 4.4). Die Verknüpfung des Tags mit dem Wikipedia-Artikel hat zur Folge, dass die Bedeutung des Tags auf den Inhalt des Wikipedia-Artikels bezogen ist und somit eine eindeutige semantische Unterscheidung getroffen wird. In der Wikipedia hat jeder Artikel einen eindeutigen Bezeichner, so dass die Programmiersprache als "Java (Programmier-sprache)" von "Java (Insel)" abgegrenzt wird. Wird kein Wikipedia-Artikel gefunden, der die Bedeutung des Tags beschreibt, können die Benutzer einen passenden Suchtreffer in Google als neues Konzept in Faviki definieren. Für die Ressource kann der Benutzer Titel, URL, Beschreibung und Kommentar speichern. Der Referenzlink vom Tag zum Artikel in Wikipedia oder zu einem Suchtreffer in Google ist die einzige vom Benutzer zusätzlich erfassbare Information für Tags. Weitere Informationen wie Abstract, Bilder etc. zum Tag werden automatisch bereitgestellt. Relationen zwischen Tags können nicht durch den Benutzer definiert werden. Da es in vielen Sprachen eine große Zahl an Wikipedia-Artikeln gibt, ist Faviki nicht auf die englische Sprache begrenzt. Tags und Ressourcen werden im persönlichen Gültigkeitsbereich eines Nutzers gespeichert.

⁸ <http://www.faviki.com>, Zuletzt abgerufen am 09.02.2011

The screenshot shows the Fuzzzy.com website interface. At the top, there's a navigation bar with links like Home, Login, Sign up, and Feedback. Below this is a search bar and a list of categories: Bookmarks, Books, Audio, Video, Feeds, Events, Tags, People, and Governance. The main content area is titled 'Semantic Web' and includes a description: 'the extension of the World Wide Web that enables people to share content beyond the boundaries of applications and websites.' It also mentions 'Also known as: sw, semweb.' and 'Created by roylac at 4. desember 2006'. On the left, there's a list of 'Related tags (14)' including Topic Maps, Web 3.0, Pragmatic Web, Semantic Web Application, SPARQL, sioc, online ask and answer, semantic desktop, semantic web adoption, RDF, Ontology, sociosemantic web, Project, and Semantic. On the right, there's a section 'Resources tagged with this' with filters for Bookmarks (219), Books (1), Audio (1), Video (3), Feeds (27), and Events (11). Below this, there are several resource cards, each with a thumbnail, title, description, tags, and a date added. The resources include 'JeromeDL - e-Library with Semantics - Home', 'JeromeDL', 'Semantic Web Client Library', 'Happy New Year: What's Ahead for the Semantic Web (Part 1)', and 'Semantic Web Site Hopes To Break Down Trade Barriers'.

Abbildung 4.5: Screenshot von Fuzzzy

Fuzzzy⁹ verwaltet die Tags im Gegensatz zu den o.g. Tagging-Systemen nicht im Common-Tag-Format, sondern auf der Basis von Topic Maps. Abbildung 4.5 zeigt einen Screenshot zum Tag "Semantic Web". Ressourcen werden in Fuzzzy in Bookmarks (Hyperlinks), Bücher (Artikel/ elektronische Bücher), Audio, Video, Newsfeeds und Ereignisse unterteilt. Für eine Ressource können Titel, URL und Beschreibung gespeichert werden. Des Weiteren kann die Angabe des inhaltlichen Stils (z.B. akademisch, lustig) und die Beschreibungstiefe (z.B. bietet Überblick, enthält detaillierte Informationen) hinzugefügt werden. Zusätzliche Informationen über Tags können durch eine Beschreibung hinzugefügt werden. Zwischen Webseite und Tag sind folgende Relationen möglich: "handelt von", "hat Beitragenden", "erwähnt", "ist eine Art von", "entspricht" und "hat". Auch Relationen zwischen Tags sind möglich, d.h. es können durch die Nutzer Synonymrelationen (siehe Screenshot "Semantic Web - Also known as: sw, semweb"), Verwandtschaftsrelationen und Ober-/Unterbegriffsrelationen erstellt werden, um die Tags semantisch genauer zu beschreiben. Tags und Webseiten besitzen globale Gültigkeit für die gesamte Community. Für

⁹ <http://www.fuzzzy.com/>, Zuletzt abgerufen am 09.02.2011

die Ressource und deren Tags wird daher der Nutzer vermerkt, der sie hinzugefügt hat. Die Nutzer von Fuzzzy können für oder gegen bereits vergebene Tags oder definierte Relationen stimmen.

4.2 Facettiertes Tagging

Eine weitere Form des semantischen Taggings ist facettiertes Tagging. Eine Facette ist eine Dimension, mit der eine Perspektive eines Objektes beschrieben werden kann. Beispielsweise können Weine anhand ihrer Anbaugebiete, Rebsorten, Geschmacksprofile (trocken, lieblich) oder auch zu welchem Essen sie passen (Geflügel, Fisch) sortiert werden. Wird ein Wein mehreren dieser Dimensionen zugeordnet, kann er über jede dieser Dimensionen gefunden werden, und des Weiteren kann die Suche über diese Dimensionen eingegrenzt werden (z.B. lieblich *und* passt zu Geflügel). Man spricht von facettnavigation, wenn eine Webseite die Navigation mittels Facetten ermöglicht, Beispiele hierfür sind Amazon¹⁰ (Facetten wie z.B. Kategorie, Format (Taschenbuch), Schriftgröße oder Preis) und Ebay¹¹ (Facetten wie z.B. Kategorie, Verpackungsart, Angebotsformat, Ort). Ein Überblick über den Einsatz von Facetten beim Social Tagging wird in [111] gegeben. Im Folgenden werden zwei Anwendungen beschrieben, die eine facettnavigation mit Tags bzw. facettnavigation ermöglichen sollen.

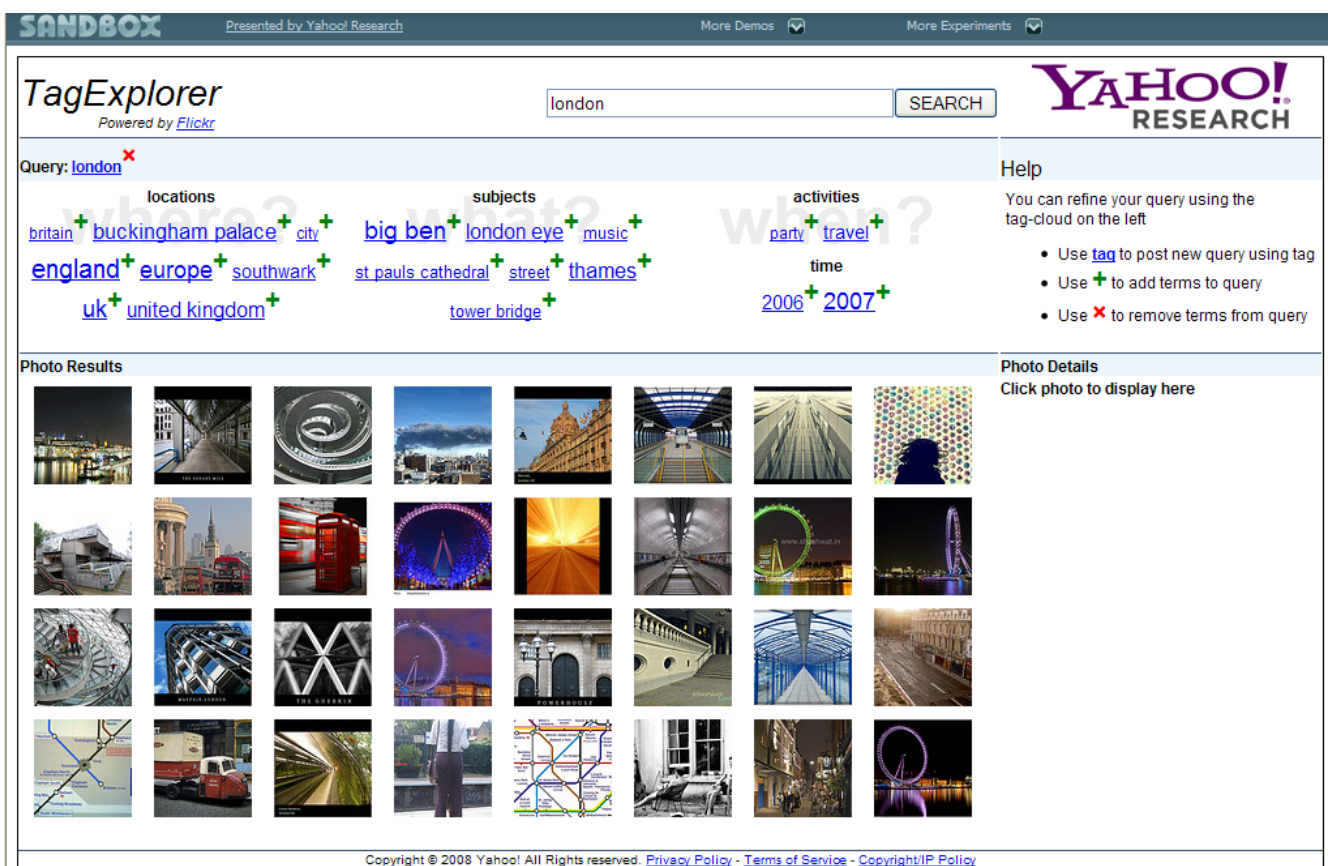


Abbildung 4.6: Facettierte Navigation im Tag Explorer

¹⁰ <http://www.amazon.com>, Zuletzt abgerufen am 27.02.2011

¹¹ <http://www.ebay.com>, Zuletzt abgerufen am 27.02.2011

Der TagExplorer¹² [140] realisiert facetiertes Tagging für Bilder in Flickr. Er wird von Yahoo entwickelt und ist seit ca. Mai 2009 als Demo veröffentlicht. Tags aus Flickr werden automatisch zu den folgenden Facetten zugeordnet:

- **Wo:** Orte, an denen das Bild aufgenommen wurde.
- **Was:** Dinge (Tiere, Gebäude, Pflanzen) und Personen, die auf dem Bild zu sehen sind.
- **Wann:** Aktivitäten bzw. Ereignisse, bei denen das Bild aufgenommen wurde und der Zeitpunkt der Aufnahme.

Abbildung 4.6 zeigt die Suche nach Bildern in London. In den Facetten werden weitere Tags eingeblendet, die gemeinsam mit "London" für Bilder vergeben wurden. Beispielsweise existieren Bilder zum Tag "London", zu denen von den Flickr-Nutzern auch Tags der Was-Facette (big ben, tower bridge) oder auch Wann-Facette (2006, 2007) vergeben wurde. So wird beispielsweise eine Navigation nach Bildern ermöglicht, die in London die Tower Bridge zum Zeitpunkt 2007 zeigen.

Für Delicious wird die Anwendung Facette¹³ von der Haystack Group des Massachusetts Institute of Technology (MIT)¹⁴ entwickelt und zum Zeitpunkt der Verfassung dieser Arbeit ersten Studien mit Endnutzern unterzogen.

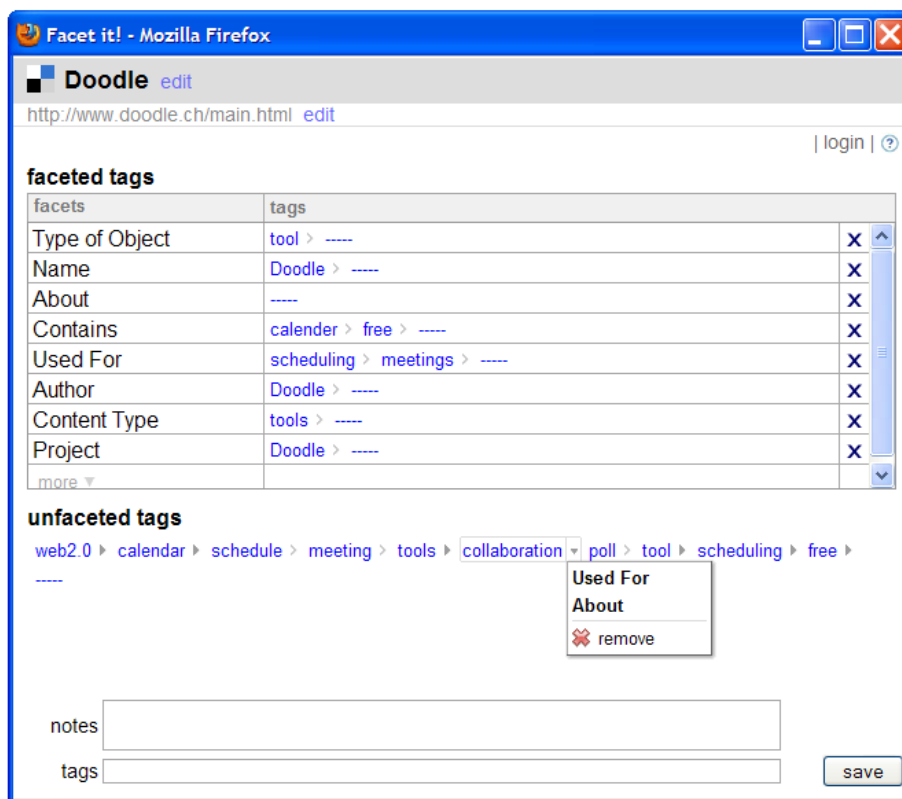


Abbildung 4.7: Tagging Dialog in Facette

In Abbildung 4.7 ist der Tagging-Dialog von Facette abgebildet. Es existieren Facetten wie Typ des Objekts (Type of Object), Name, Thema (About), "Enthält" (Contains), Nutzungskontext (Used For) und

¹² <http://tagexplorer.sandbox.yahoo.com>, Zuletzt abgerufen am 27.02.2011

¹³ <http://facette.csail.mit.edu>, Zuletzt abgerufen am 27.02.2011

¹⁴ <http://groups.csail.mit.edu/haystack/>, Zuletzt abgerufen am 27.02.2011

Autor (Author). Beim Taggen kann der Nutzer ein Tag in der entsprechenden Zeile des Formulars eintippen, oder frei im Bereich der unfacettierten Tags schreiben und später einer Facette zuordnen. Wie in der Abbildung gut zu sehen ist, kann ein Tag gleichzeitig in verschiedene Facetten eingeordnet werden, beispielsweise Doodle als Name und Projekt.

The screenshot shows the 'Facette Delicious Browser' interface. On the left, there are several faceted navigation panels: 'Type of Object' (100 items), 'About' (80 items), 'Contains' (72 items), 'Location' (1 item), 'Written In' (2 items), 'For' (6 items), 'Name' (40 items), 'Author' (27 items), 'Source' (23 items), 'Used For' (21 items), and 'Update Frequency' (7 items). The main content area displays a list of 100 bookmarks, sorted by 'add date'. The first 10 items are visible, including 'Home | Scribd', 'Doodle', 'Free! Why \$0.00 Is the Future of Business', 'Jester: The Online Joke Recommender', 'My Yahoo!', 'Moserware: Does Your Code Pass The Turkey Test?', 'Google ready to invade telco turf, bid on 700MHz spectrum', 'New Turbine Design Wins Sustainability Award : TreeHugger', 'Pandora Radio - Listen to Free Internet Radio, Find New Music', and 'Official Google Blog: An update on Google Video'. On the right, there is a 'Tags' panel showing a list of tags and their counts, such as 'article (34)', 'blog (23)', 'tips (14)', 'tool (13)', 'art (12)', 'news (12)', 'culture (11)', 'programming (11)', 'technology (11)', 'design (10)', 'howto (10)', 'reference (10)', 'homepage (8)', 'humor (8)', 'inspiration (8)', 'business (7)', 'daily (7)', 'food (7)', 'web2.0 (7)', 'cooking (6)', 'recipes (6)', 'tutorial (6)', 'web_design (6)', 'education (5)', 'lifehacks (5)', 'magazine (5)', 'reviews (5)', 'software (5)', 'graphics (4)', 'green (4)', 'politics (4)', 'productivity (4)', 'psychology (4)', 'shopping (4)', 'web_development (4)', 'activism (3)', 'advice (3)', 'baking (3)', 'blog_post (3)', 'css (3)', 'diy (3)', 'drawings (3)', and 'enviornment (3)'.

Abbildung 4.8: Facettierte Navigation in Facette

Abbildung 4.8 zeigt einen Ausschnitt aller Facetten am linken Rand, mit denen die Navigation der Bookmarks von Delicious in Facette eingegrenzt werden kann.

4.3 Typisiertes Tagging

Eine andere Form des semantischen Taggings ist typisiertes Tagging, bei dem statt des Links zu einem beschreibenden Konzept in einer externen Datenquelle der Typ des Tags angegeben wird. Typisiertes Tagging ist eng verwandt mit facettiertem Tagging. Facettiertes Tagging ordnet die Tags abhängig von der Ressource in verschiedene Dimensionen ein. Typisiertes Tagging, wie es in dieser Arbeit definiert wird, klassifiziert den Typ eines Tags unabhängig von der Ressource. Beispielsweise ist das Tag "Ralf Steinmetz" eine Person, unabhängig davon, ob die Ressource von "Ralf Steinmetz" handelt und demzufolge eher einer thematischen Facette zugeordnet werden müsste, oder ob "Ralf Steinmetz" der Autor der Ressource ist.

4.3.1 Definition der Basismenge der Tag-Typen für Ressourcen-basiertes Lernen

In der Literatur sind durch die Analyse von Tags verschiedene Arten/Typen von Tags bekannt, siehe Kapitel 3.1.3. Diese Typen werden in diesem Abschnitt hinsichtlich ihrer Verwendung im Ressourcen-basierten Lernen untersucht, um eine Basismenge von Tag-Typen zu definieren. Die Basismenge soll das Anwendungsszenario abdecken, muss aber gleichzeitig überschaubar bleiben, damit sie intuitiv benutzbar ist. Tabelle 4.1 fasst identifizierte Typen aus ausgewählter Literatur zusammen und stellt sie den in dieser Arbeit definierten Tag-Typen gegenüber. In der vorletzten Zeile der Tabelle ist angegeben, welche Tagging-Systeme in der Arbeit betrachtet werden. Die letzte Zeile gibt Auskunft darüber, ob in den Arbeiten ein Verfahren zur automatischen Erkennung dieser Tag-Typen beschrieben wird.

Golder und Humberman [47] (siehe auch Kapitel 3.1.3) haben Tags in Delicious analysiert und die folgenden Tag-Typen identifiziert:

- **Thema** (Identifying what (or who) it is about): Diese Tags identifizieren, von was oder wem das Objekt handelt.
- **Typ** (Identifying what it is): Diese Tags identifizieren den Typ des Objekts, z.B. "Artikel", "Blog", "Buch".
- **Autor oder Besitzer** (Identifying who owns it): Mit diesen Tags wird der Ersteller oder der Besitzer des Objektes identifiziert.
- **Verfeinerungskategorien** (Refining categories): Dies sind individuelle Tags, die für sich allein genommen keinen Sinn ergeben, sondern die genutzt werden, um existierende Kategorien zu verfeinern (z.B. Tags wie "25" und "100" erfüllen die Funktion, eine Kategorie weiter zu unterteilen um Unterkategorien zu bilden).
- **Bewertungen** (Identifying qualities or characteristics): Meinungen des Nutzers werden in dieser Funktionsgruppe ausgedrückt, z.B. "lustig" oder "interessant".
- **Selbstreferenz** (Self reference): Diese Tags dienen dazu, die Relation zwischen Objekt und Nutzer auszudrücken. Häufig beginnt das Tag mit "my", z.B. "mystuff" oder "mycomments".
- **Aufgabenmanagement** (Task organizing): Beim Sammeln von Objekten für eine bestimmte Aufgabe werden Tags zum Aufgabenmanagement genutzt, beispielsweise die Tags "toRead" oder "job-search".

Golder und Huberman [47]	Sauermann [124]	Bischoff et al. [16]	Wartena [156]	TagExplorer [140]	Eigener Ansatz
Thema	Thema	Thema	Thema	Gegenstand	Thema
Typ	Medientyp	Typ	Attribut		Typ
Autor/Besitzer	Person, Organisation	Autor/Besitzer	Autor	Person/Gruppe	Person/Organisation
Bewertung		Bewertung	Meinung		
Aufgabe		Nutzungskontext			Ziel
Selbstreferenz		Selbstreferenz	Selbstreferenz		
	Ort	Ort		Ort	Ort
	Prozess	Zeit		Aktivität/Ereignis Zeit	Ereignis
Verfeinerungskategorien					
Delicious	Gnowsis	Delicious, Flickr, Last.fm	LibraryThing	Flickr	ELWMS.KOM
		Automatische Erkennung der Tag-Typen	Automatische Erkennung der Tag-Typen	Automatische Erkennung der Tag-Typen	Automatische Erkennung der Tag-Typen

Tabelle 4.1: Verwendete Tag-Typen in verwandten Arbeiten

Sauermann [124] hat ein Modell für persönliches Informationsmanagement (siehe auch Kapitel 6.2) bestehend aus den folgenden Basistypen für sein Anwendungssystem "Gnowsis" entwickelt:

- **Ort** (Location): umfasst Orte, für die er weitere Unterklassen (Gebäude, Stadt, Land,...) modelliert hat.
- **Medientyp** (LogicalMediaType): logische Medientypen, z.B. "Buch", "Vertrag", "Werbevideo", "To-do-Liste", "Flyer", "Angebot", "Report", "Hausaufgabe" ...
- **Organisation** (Organization): Verwaltungs- bzw. Funktionsstrukturen, z.B. Unternehmen, politische Parteien ...
- **Person**: Person (gleichgültig ob lebendig, tot, real oder erfunden)
- **Prozess** (ProcessConcept): Konzepte, die sich auf eine Serie von Aktionen oder Tätigkeiten beziehen, die zu einem Ende führen, z.B. Ereignis, Meeting, Projekt und Aufgabe
- **Thema** (Topic): Gegenstand einer Diskussion oder eines Dokuments, beispielsweise wissenschaftliche Fachgebiete.

Bischoff et al. [16] haben Tags in Delicious, Flickr und Last.fm¹⁵ (Musik) analysiert. Im Folgenden wird hauptsächlich die Analyse der Tags in Delicious betrachtet, da diese dem in dieser Arbeit adressierten Anwendungsszenario am ehesten entsprechen. Dort werden die folgenden 8 Tag-Typen von Bischoff et al. unterschieden:

- **Zeit** (Time): Datumsangaben, Feiertage, "daily", "current", "2004" ...
- **Ort** (Location): Städte, Länder ...
- **Typ/Medium** (Type): Dateiformate und Webgenres wie Blog und Wiki ...
- **Autor/Besitzer** (Author/Owner)
- **Selbstreferenz** (Self reference): "myfriends", "wishlist" ...
- **Thema** (Topic)
- **Meinung** (Opinion): "interesting", "relevant" ...
- **Nutzungskontext** (Usage context): "toRead", "review", "later" ...

Wartena [156] hat Tags im Tagging-System LibraryThing¹⁶ analysiert. In LibraryThing können Bücher mittels Tagging verwaltet werden. Seine Arbeit identifiziert die folgenden Tag-Typen:

- **Thema** (Topic): thematischer Inhalt des Buches
- **Autor** (Author): Autor des Buches
- **Attribut** (Attribute): umfasst Genre ("Historische Fiktion", "Schwedische Literatur") und Nutzungskontext bezogen auf das literarische Werk ("UK Edition", "Übersetzung")
- **Selbstreferenz** (Self reference): umfasst organisatorische Tags, die sich auf ein bestimmtes Exemplar des Buches beziehen (z.B. "ausgeliehen" und "Wasserschaden")
- **Meinung** (Opinion): die Meinung des Lesers

Thema, Autor und Attribut fasst er zusammen als Klasse "Werk", da sie sich auf das Buch beziehen und Selbstreferenz und Meinung sind auf den Nutzer bzw. Besitzer des getaggten Buches bezogen.

¹⁵ <http://www.lastfm.de/>, Zuletzt abgerufen am 02.03.2011

¹⁶ <http://www.librarything.com/>, Zuletzt abgerufen am 02.03.2011

Der TagExplorer [140] klassifiziert Flickr-Tags in den folgenden Kategorien:

- **Ort** (Location)
- **Gegenstand** (Subject): z.B. Tier, Pflanze und Gebäude
- **Person und Gruppe von Personen** (Names)
- **Aktivität** (Activity) z.B. "Triathlon",... bzw. **Ereignis** (Event) z.B. "Valentines day",...
- **Zeit** (Time) z.B. "august", "2006",...

Da im adressierten Anwendungsszenario meist Ressourcen fremder Autoren getaggt werden, gibt es wenig Möglichkeiten für Selbstreferenzen. Deshalb wird diese Kategorie nicht in die Basismenge der Tag-Typen übernommen. Verfeinerungskategorien können besser durch Relationen der Tags untereinander ausgedrückt werden als durch eigene Kategorien, daher hat diese Kategorie keine Relevanz für die zu definierende Basismenge. Die Kategorien Person und Organisation können auf Grund ihrer starken Ähnlichkeit in der Basismenge zusammengefasst werden. Ein Ziel beschreibt den Nutzungskontext der Ressource und ist entsprechend der Anforderungsanalyse (siehe Kapitel 7.1) ein sehr wichtiger Tag-Typ, der in die Basismenge aufgenommen wird. Ressourcen werden mit Zielen getaggt und der Beitrag der Ressource zur Zielerreichung (z.B. Relevanz und Qualität der Ressource) kann vom Nutzer mittels einer entsprechenden Funktion angegeben werden. Daher bildet die Kategorie der Meinungs-Tags keinen eigenen Tag-Typ im adressierten Anwendungsszenario. Tags, die keinem Tag-Typ zugeordnet werden können, sind generell möglich, werden aber nicht typisiert.

Damit wurde die nachfolgende Basismenge für das Anwendungsszenario des Ressourcen-basierten Lernens definiert:

- **Thema** - z.B. von welchem Thema handelt die Ressource, ...
- **Typ** - z.B. was ist das Webgenre der Ressource (z.B. Blog, Wiki, Wissenschaftliche Publikation), was ist der Dateityp (z.B. PDF),...
- **Person/Organisation** - wer hat die Ressource veröffentlicht, wer hat sie empfohlen, mit wem verbinde ich die Ressource, wer hat sie präsentiert, ...
- **Ziel** - beispielsweise in welchem Zusammenhang wurde die Ressource gefunden (z.B. Erstellen eines Vortrags XY),...
- **Ereignis** - z.B. welches Ereignis steht im Zusammenhang mit der Ressource (z.B. wurde empfohlen während der Konferenz YZ),...
- **Ort** - wo wurde die Ressource veröffentlicht/präsentiert oder wo wurde die Ressource gefunden (z.B. während Gastaufenthalt in AB wurde die Ressource recherchiert), ...

4.3.2 Evaluation der definierten Basismenge

Die definierte Basismenge der Tag-Typen wurde nachfolgend in mehreren Benutzerstudien hinsichtlich ihrer Vollständigkeit überprüft. Dabei wurden einerseits die Nutzer befragt und andererseits die tatsächliche Nutzung der Tag-Typen gemessen.

Abbildung 4.9 schlüsselt den Anteil der verwendeten Tag-Typen an allen Tags je Benutzerstudie auf. Details zum Ablauf der Evaluationen finden sich in Kapitel 9. Dabei ist zu beachten, dass in der Tabelle nicht die einzelnen Tag-Instanzen aufgelistet sind. Damit ist gemeint, dass ein Tag, das für 10 Ressourcen

vergeben wurde, nicht 10-mal gezählt wird, sondern nur einmal. Es wurden während der Studien insgesamt 1611 Tags erstellt. Insgesamt wurden in allen Studien alle Tag-Typen genutzt, wenn auch nicht jeder Tag-Typ von jedem Teilnehmer.

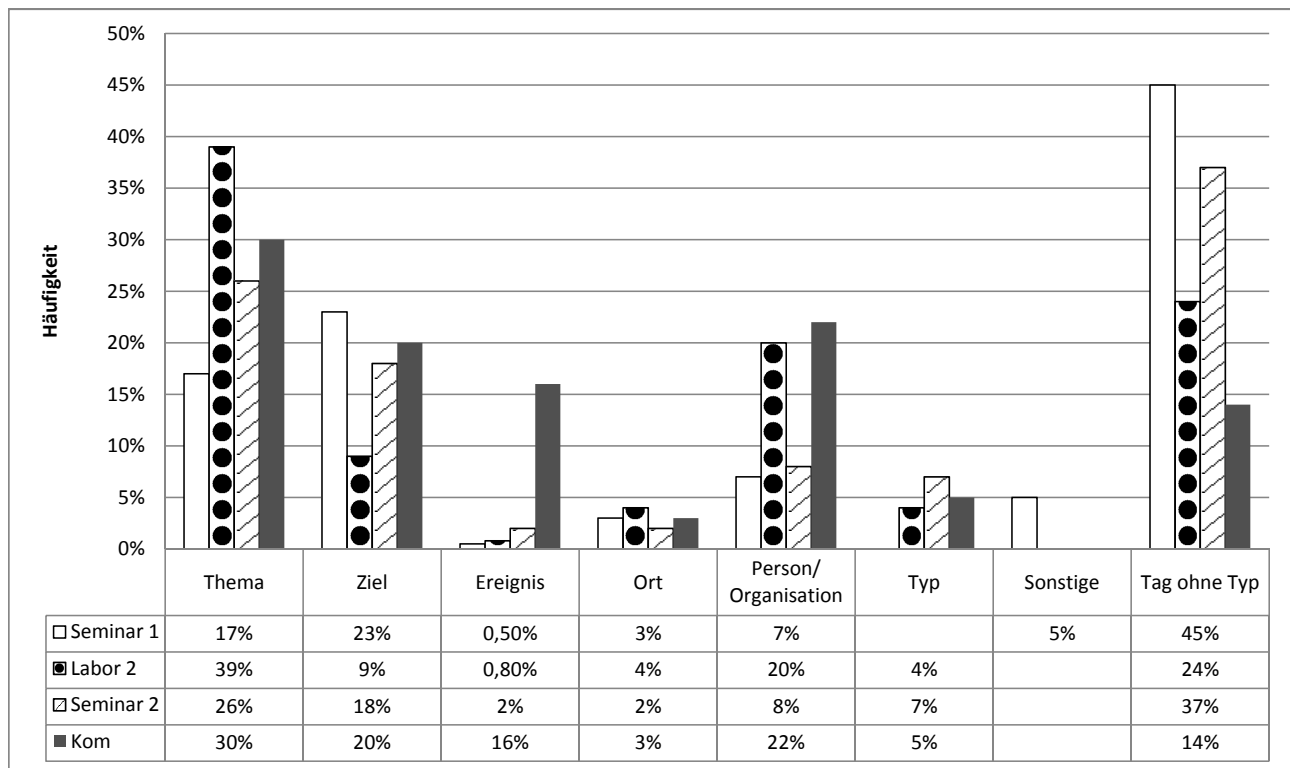


Abbildung 4.9: Anzahl der typisierten Tags in Evaluationen

In der Studie "Seminar 1" traten Tags, denen kein Tag-Typ zugewiesen wurde, mit 45% am häufigsten auf. Durch die Umgestaltung des Tagging-Dialogs (siehe Abbildungen 8.5 und 8.6 in Kapitel 8.2) konnte der Anteil der untypisierten Tags erheblich reduziert werden. Der Tag-Typ "Sonstiges" in der Studie "Seminar 1" diente dazu, eventuelle weitere, in der Festlegung der Basismenge nicht enthaltene Typen zu identifizieren. Es zeigte sich nach qualitativer Rückmeldung einiger Nutzer und manueller Durchsicht der untypisierten Tags, dass der Tag-Typ "Typ" wichtig ist, so dass ab der Studie "Labor 2" der Tag-Typ "Typ" eingeführt wurde. Dies wird auch anhand der Tags in Delicious ersichtlich, denn Tags wie "blog", "wiki" und "forum", die zum Tag-Typ "Typ" zugeordnet werden können, tauchen in der Tag Cloud der meist verwendeten Tags auf (siehe Abbildung 6.5). Die in allen Studien verwendete Zielmanagementfunktion führte dazu, dass im Verhältnis sehr viele Tags vom Typ Ziel erstellt wurden.

Die Teilnehmer der Studie "KOM" wurden befragt, welche der Tag-Typen sie für notwendig erachten. Die Ergebnisse sind in Tabelle 4.2 aufgelistet. Die Werte decken sich in Relation mit den tatsächlich verwendeten Tag-Typen in Tabelle 4.9: Themen wurden mit 30% Häufigkeit am meisten genutzt und werden von allen Teilnehmern gewünscht. Die zweithäufigsten Typen sind Ziel (20%) und Person/Organisation (22%), die auch in der Befragung von 72% und 78% der Teilnehmer gewünscht werden. Ereignisse, Ort und Typ werden seltener verwendet, sind aber für jeweils mindestens 50% der Teilnehmer wichtig.

Thema	Ziel	Ereignis	Ort	Person/ Organi- sation	Typ
100%	70%	50%	55%	75%	50%

Tabelle 4.2: Bedarf der Tag-Typen bei Probanden

Neben der Frage nach der Wichtigkeit der Tag-Typen wurden die Probanden befragt, welche Tag-Typen sie sich außerdem wünschen. 50% der Teilnehmer sind mit den vorhandenen Tag-Typen zufrieden. Die Antworten der anderen Hälfte der Teilnehmer sind wie folgt:

- "for:XY" Tag wie in Delicious, mit dem man gezielt anderen Nutzern Ressourcen empfehlen kann (1 Nennung)
- Qualität (1 Nennung)
- Sprache (1 Nennung)
- Aktualität (1 Nennung)
- Organisation | Gruppe | Firma (2 Nennungen)
- Jahr (2 Nennungen)

Hierbei ist anzumerken, dass Konzept und Realisierung des Werkzeugs (siehe Kapitel 7) die Erweiterungen der Basistagtypen durch die Nutzer vorsieht und ermöglicht.

4.3.3 Nutzen der typisierten Tags

Das Typisieren von Tags ermöglicht eine Navigation ähnlich dem facettierten Navigieren, d.h. die Menge der Ressourcen kann anhand der Tag-Typen in verschiedenen Dimensionen eingegrenzt werden. Dadurch werden beispielsweise Anfragen möglich wie "Zeige alle Ressourcen, die mit dem Ort 'Darmstadt' und der Person 'Ralf Steinmetz' getaggt wurden" ohne eine spezielle Suchsyntax zu verwenden.

Des Weiteren können spezielle Ansichten für die einzelnen Tag-Typen realisiert werden, beispielsweise können Ressourcen anhand ihres Orts-Tag auf einer geographischen Karte angezeigt werden oder anhand eines Ereignisses in einer Kalenderansicht. Auch das Zurechtfinden in der Tagging-Struktur von anderen Nutzern kann durch die Tag-Typen erleichtert werden.

Ein weiterer Nutzen ist die Verbesserung von Algorithmen zur Empfehlung von Tags und Ressourcen. Dieser Aspekt wird im folgenden Abschnitt "Recommendation von Tags und Ressourcen" näher erläutert.

4.4 Recommendation von Tags und Ressourcen

Zu sogenannten Folksonomien (siehe Kapitel 3.2.6) aggregierte Tags sind häufig Grundlage für das Vorschlagen von Tags und Ressourcen, sogenannte Recommendations. Im folgenden wird knapp in die Thematik eingeführt, für einen Überblick über Recommender-Verfahren wird auf die Literatur, z.B. [107], verwiesen. Des Weiteren werden in dieser Arbeit realisierte Tag- und Ressourcen-Empfehlungen erläutert, die durch die typisierten Tags möglich werden.

4.4.1 Recommendation von Tags

Unter Tag Recommendation wird das Vorschlagen von Tags in Tagging-Systemen verstanden. Dabei gibt es verschiedene Möglichkeiten, Tags für das Vorschlagen auszuwählen, u.a.:

- **Inhaltsbasierte Methoden:** Diese Methoden analysieren den Inhalt der zu taggenden Ressource, im Falle von Webseiten demzufolge den Inhalt der Webseite. Dies bedeutet, dass der Inhalt der Webseite analysiert wird und beispielsweise wichtige Stichwörter im Text extrahiert werden, die dem Nutzer als Tags vorgeschlagen werden.
- **Ressourcenbasierte Methoden:** Diese Methoden analysieren die Tags, die von anderen Nutzern für die zu taggende Ressource vergeben wurden und schlagen beispielsweise häufig verwendete Tags vor. Des Weiteren dehnen manche der Methoden die Analyse auf weitere Ressourcen aus, die mit Tags verschlagwortet sind, mit denen auch die zu taggende Ressource bereits getaggt wurde.
- **Nutzerbasierte Methoden:** Diese Methoden basieren auf der Tag-Historie des taggenden Nutzers und schlagen Tags vor, die ein Nutzer bereits verwendet hat. Des Weiteren gibt es Methoden, die den aktuellen Aufenthaltsort des Nutzers und den Zeitpunkt des Taggings einbeziehen, z.B. um Tags für Fotos vorzuschlagen.
- **Hybride Methoden:** Diese Methoden kombinieren verschiedene Methoden miteinander.

Da Nutzer auch Tags verwenden, die lediglich Selbstreferenzen darstellen oder der persönlichen Organisation der Ressourcen dienen, und daher keine nutzerübergreifende Bedeutung haben, kann dies Empfehlungen beeinträchtigen. Somit können Verfahren, die Tags anderer Nutzer einbeziehen, davon profitieren, dass Tag-Typen, die nur eine individuelle Bedeutung für den taggenden Nutzer haben wie "toRead" oder "funny", von dem Verfahren nicht berücksichtigt werden. Tags wiederum, die als Thema typisiert wurden, können von dem nachfolgend beschriebenen Verfahren stärker gewichtet werden, da Themen nutzerübergreifend von Interesse sind und eine gewisse Aussagekraft über die getaggte Ressource haben.

Im implementierten System (siehe Kapitel 8) wurde eine Kombination aus ressourcen- und nutzerbasierter Methode umgesetzt. Zum einen werden Tags vorgeschlagen, die auch andere Nutzer für die zu taggende Webseite bereits vergeben haben. Dabei werden nur Themen, Personen, Orte und Ereignisse berücksichtigt. Zum anderen werden Suchhistorie, Tagging-Historie und Ziele des Nutzers in die Empfehlungen einbezogen. Suchbegriffe, die der Nutzer vor dem Speichern der Webseite eingegeben hat, um die Webseite zu finden, werden hierbei als Tags für die Webseite vorgeschlagen. Tags, die der Nutzer während einer Recherche bereits für frühere Ressourcen verwendet hat, werden ebenfalls als Tags vorgeschlagen. Da Ressourcen im Anwendungsszenario vor allem auch Zielen zugeordnet werden sollen, gibt es für die Nutzer die Möglichkeit im System, ein Ziel als aktuell verfolgtes Ziel zu aktivieren. Dieses aktivierte Ziel wird als Tag für die zu speichernde Ressource vorgeschlagen, bis das Ziel vom Nutzer deaktiviert wird.

Beim Vorschlagen der Tags wird deren Tag-Typ angezeigt, um die Übersicht über die vorgeschlagenen Tags zu erhöhen, siehe Screenshot in Abbildung 8.6.

Evaluation

Die implementierte Methode zur Empfehlung von Tags wurde in den Studien "Labor 2", "Seminar 2" und "KOM" mit Benutzern evaluiert (siehe auch Kapitel 9). Tabelle 4.3 listet die durch die Benutzer anteilige Verwendung der verschiedenen vorgeschlagenen Tags auf. In der letzten Zeile der Tabelle ist der Gesamtanteil der vorgeschlagenen und akzeptierten Tags in Bezug zu allen tatsächlich vergebenen Tags aufgeführt.

	Labor 2	Seminar 2	KOM
Autovervollständigung	20%	24%	15%
Vorschläge aus Wissensnetz	3%	2%	2%
Vorschläge aus Suchwort	6%	4%	7%
Vorschläge zuletzt verwendete Tags	20%	26%	14%
Aktiviertes Ziel	16%	24%	28%
Gesamtanteil der verwendeten Tags	65%	80%	66%

Tabelle 4.3: Ergebnisse der Evaluation der Tag Recommendation (Tatsächliche Verwendung)

Neben den Tag-Vorschlägen wurde auch eine Autovervollständigungs-Funktion implementiert, die nach Tippen der ersten Buchstaben entsprechende, im Wissensnetz bereits existierende Tags auflistet. Zwischen 15% und 24% aller verwendeten Tags wurden mittels dieser Funktion vergeben. Nur zwischen 2% bis 3% der verwendeten Tags konnten auf Basis bereits gespeicherter Ressourcen erfolgreich empfohlen werden. Erfolgreich bedeutet hier, dass der Nutzer die Empfehlung angenommen hat. Ressourcen-basierte Methoden benötigen in der Regel eine kritische Masse an Nutzern bzw. getaggten Ressourcen. Diese kritische Masse konnte während der Evaluation nicht erreicht werden, wie die Zahlen gut ausdrücken. Insgesamt wurden auf Basis dieser beiden Methoden weniger als ca. ein Viertel der verwendeten Tags erfolgreich empfohlen. Diese Werte konnten durch den Einsatz nutzerbasierter Methoden auf 65% bzw. 80% deutlich erhöht werden, was einen geringeren manuellen Aufwand beim Tagging für die Nutzer zur Folge hat.

Vorgeschlagene und akzeptierte Suchbegriffe haben nur einen Anteil von 4% bis 7% der verwendeten Tags. Ausgewertet wurden hier allerdings nur Suchbegriffe, die in der Suchmaschine von Google eingegeben wurden, da diese die am häufigsten verwendete Suchmaschine unter den Studienteilnehmern ist. Jeder Suchbegriff wird dabei als eigenes Tag verwendet, Leerzeichen dienen somit zur Trennung zwischen den vorgeschlagenen Tags, außer die Suchbegriffe wurden in Anführungsstrichen eingefasst. Es ist zu vermuten, aber nicht bewiesen, dass die Suchbegriffe, die aus mehreren Wörtern bestehen, häufig nicht in Anführungsstrichen eingegeben werden und die Suchbegriffe nicht als einzelne Tags taugen. Tags, die aus der Menge der zuletzt verwendeten Tags vorgeschlagen wurden, bilden mit 14% bis 26% einen großen Anteil. Auch die Funktion zur Aktivierung eines Ziels ist mit einem Anteil von 16% bis 28% erfolgreich.

Die Teilnehmer der Studie "KOM" wurden mittels Fragebogen zur Qualität der vorgeschlagenen Tags befragt. 87% der Studienteilnehmer, die sich bei dieser Frage nicht enthalten haben, stimmen (eher) zu, dass die Tags, die beim Tagging vorgeschlagen wurden, das Tagging gut unterstützt haben.

4.4.2 Recommendation von Ressourcen

In Tagging-Systemen können den Nutzern nicht nur Tags vorgeschlagen werden, sondern es können ihnen auch Ressourcen empfohlen werden. Methoden für das Empfehlen von Ressourcen analysieren beispielsweise die Folksonomie und können so auf Basis der getaggtten Ressourcen und auf Basis des Nutzerverhaltens in der Vergangenheit dem Nutzer Ressourcen empfehlen.

Auch die Verfahren zur Empfehlung von Ressourcen, die die Tags der Nutzer einbeziehen, können davon profitieren, dass Tags, die nur eine individuelle Bedeutung für den taggenden Nutzer haben wie "toRead" oder "funny", von dem Verfahren nicht berücksichtigt werden. Das Tag "toRead" weist z.B. auf keine Ähnlichkeit der Ressourcen hin. Tags wiederum, die als Thema typisiert wurden, können vom Verfahren stärker gewichtet werden, da Themen nutzerübergreifend von Interesse sind und eine gewisse Aussagekraft über die getaggte Ressource haben.

Des Weiteren werden in bekannten Systemen nur Tags analysiert, die direkt mit der Ressource verbunden sind, da im Allgemeinen in Tagging-Systemen keine Relationen zwischen Tags existieren. Dies ist in Tagging-Systemen mit einer großen Masse an Nutzern ausreichend, allerdings funktioniert das Verfahren zur Recommendation von Ressourcen bei neu gespeicherten Ressourcen nur schlecht, da diese noch spärlich getaggt sind. Im in dieser Arbeit vorgeschlagenen Tagging-System sind Relationen zwischen Tags möglich (siehe Kapitel 7), die vom Recommendation-Verfahren genutzt werden können, d.h. das Verfahren kann die Struktur der Tags einbeziehen. Abbildung 4.10 zeigt ein Beispiel für eine solche strukturelle Recommendation. Es wurden die zwei Ereignisse (Kalendersymbol) "AH" und "UM" mit dem gleichen Thema (Glühlampensymbol) "User Modelling for Personalization" getaggt. Zwei Ressourcen wurde mit je einem dieser beiden Ereignisse getaggt. Diese Ressourcen sind durch das gleiche Thema der Ereignisse potentiell ähnlich, auch wenn sie nicht direkt mit dem gleichen Thema getaggt wurden. Der längere Strukturpfad kann für Empfehlungen genutzt werden, wenn auf Basis der direkten Tags keine Empfehlungen erfolgen können.

Evaluation

Im im Rahmen dieser Arbeit entwickelten System ELWMS.KOM (E-Learning Wissensmanagement-System) wurden Ressourcen anhand struktureller Informationen empfohlen und durch Benutzer evaluiert (siehe auch Kapitel 9). Für die Bewertung der Nutzer liegen nur Fragebogendaten vor, da vom System die besuchten Webseiten nicht vollständig aufgezeichnet wurden, um die Privatheit der Nutzer, die nicht anonym an der Studie teilgenommen haben, zu schützen.

63% der Studienteilnehmer, die sich bei dieser Frage nicht enthalten haben, gaben an, dass die Empfehlungen geholfen haben, weitere relevante Informationen zu finden.

4.4.3 Fazit

In diesem Abschnitt wurden im Rahmen der Arbeit umgesetzte Methoden für das Vorschlagen von Tags und Ressourcen erläutert und deren Evaluation durch Benutzer beschrieben. Auf Grund der fehlenden kritischen Masse an Nutzern werden verstärkt nutzerbasierte Methoden verwendet, die die Menge der

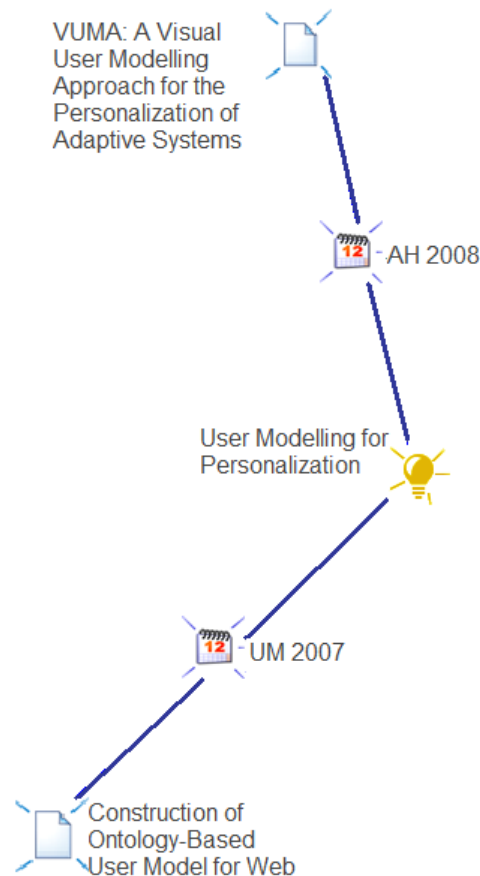



Abbildung 4.10: Strukturelle Recommendation von Ressourcen

erfolgreich empfohlenen Tags erhöhen können. Die Benutzer des Systems finden die Vorschläge überwiegend gut. Dennoch sind die Verfahren ausbaufähig. Die Kombination mit inhaltsbasierten Methoden, wie sie mittels Extended Explicit Semantic Analysis (XESA) [131] realisiert werden könnten, kann die Empfehlungen weiter verbessern, indem beispielsweise Kategorien in Wikipedia als Tags vorgeschlagen werden, deren zugeordnete Artikel ähnlich zur taggenden Ressource sind. Des Weiteren kann das Feststellen von Ähnlichkeiten zwischen Ressourcen (und Tags) die Genauigkeit der strukturellen Recommendation erhöhen. Die Ergebnisse zeigen nur Trends, und müssen zukünftig mit einer größeren Anzahl an Nutzern und Ressourcen überprüft werden. Offene Punkte bei der strukturellen Recommendation sind beispielsweise noch die maximal zu berücksichtigende Pfadlänge bei strukturellen Empfehlungen und das unterschiedliche Gewichten der verschiedenen Tag-Typen.

4.5 Zusammenfassung

In diesem Kapitel wurden die Vorteile von semantischem Tagging gegenüber einfachem Tagging erläutert. Es reduziert die Probleme, die durch unkontrolliertes Vokabular und durch flache Strukturen in Tagging-Systemen verursacht werden. Semantisches Tagging mittels Beschreibung der Tags durch eine URL ist wiederum aufwändig. Daher wurde in diesem Kapitel eine weitere Form des semantischen Taggings - das typisierte Tagging - vorgestellt. Diese Form ist für das Szenario des Ressourcen-basierten Lernens nutzbar und hilfreich, wie die Studien siehe Kapitel 9 zeigen. Jedem Tag kann beim typisierten



Tagging ein Tag-Typ zugeordnet werden. Abschließend im Abschnitt 4.4 wurde der Nutzen der typisierten Tags für Verfahren zur Empfehlung von Tags und Ressourcen erläutert.

5 Automatische Erkennung der Tag-Typen

In diesem Kapitel wird untersucht, inwieweit eine automatische Erkennung der Tag-Typen realisiert werden kann. Dieses Kapitel gliedert sich neben der Motivation für ein automatisches Verfahren und der Zusammenfassung in zwei Teile. Im ersten Teil des Kapitels werden die Grundlagen der automatischen Verfahren erläutert und es werden verwandte Verfahren zur Erkennung der Tag-Typen beschrieben. Im Rahmen dieser Arbeit wurde ein eigenes automatisches Verfahren entwickelt und evaluiert. Das Verfahren und dessen Evaluationsergebnisse werden im zweiten Teil dieses Kapitels erläutert.

5.1 Motivation

Durch eine automatische Erkennung kann der Nutzer beim Tagging von Ressourcen unterstützt werden. Auf die Frage, ob sie sich eine automatische Erkennung der Tag-Typen wünschen, um den manuellen Aufwand zu reduzieren, haben 85% der Probanden der Evaluation von ELWMS.KOM diese Frage bejaht.

Denkbar sind auch Szenarien, in denen ein Nutzer beim Speichern der Ressource zuerst untypisiert taggt und erst im späteren Verlauf den Tags einen Tag-Typ zuordnet. Dabei können in einer Art Pflegeaktion die Tags automatisch zu den verschiedenen Tag-Typen zugeordnet werden, die der Nutzer noch manuell nachbearbeiten kann. In solch einem Fall ist eine auf Basis des Tag-Typs gefilterte Gruppe von Tags leichter vom Nutzer kontrollierbar, als wenn er die Gesamtmenge untypisierter Tags überblicken und manuell zuordnen muss. Beispielsweise kann ein Mensch relativ leicht in einer Menge identifizieren, ob ein Tag als Person klassifiziert wurde, das eigentlich ein Ort ist.

Eine weitere Anwendung der automatischen Erkennung kommt bei der Empfehlung von Tags zum Tragen (siehe Kapitel 4.4). Sollen untypisierte Tags für eine Ressource vorgeschlagen werden, die noch nicht als typisierter Tag im Tagging-System existieren, können auch auf diese Tags die verbesserten Verfahren zur Empfehlung von Tags angewendet werden.

Ein weiterer Anwendungsfall für die automatische Erkennung wird sichtbar, wenn die Tags vom Tag-Typ Ziel analysiert werden: Ziele entsprechen häufig thematischen Zielen. Diese Erkenntnis deckt sich mit früheren Studien in [129]. Im Kapitel 4.4 wurde erläutert, dass sich Recommendation-Verfahren verbessern lassen, indem Tags wie beispielsweise Ziele, die im Allgemeinen nur einen individuellen Zweck erfüllen, unberücksichtigt bleiben. Eine automatische Erkennung von thematischen Zieltags, die somit nicht nur einen individuellen Zweck erfüllen, kann diese damit für das Recommendation-Verfahren zum Tag-Typ Thema zuordnen, und somit können die thematischen Ziele trotz ihres eigentlichen Tag-Typs in die Empfehlungen einbezogen werden.

5.2 Grundlagen und verwandte Ansätze

Im Folgenden werden Grundlagen für eine automatische Erkennung von Tag-Typen erläutert und verwandte Ansätze vorgestellt.

5.2.1 Eigennamenerkennung

Ein der automatischen Erkennung von Tag-Typen verwandtes Forschungsgebiet ist das der Erkennung von Eigennamen.

Die Eigennamenerkennung (Named Entity Recognition) in natürlichsprachlichen Texten ist ein Teilforschungsgebiet der Informationsextraktion. Ziel der Eigennamenerkennung ist sowohl die automatische Erkennung von Namen bzw. Wörtern von speziellem Interesse als auch die Klassifizierung der Namen in Klassen wie Personen, Orte, Organisationen etc. Für die Klassifizierung der Namen werden zumeist listen- und regelbasierte Methoden und Verfahren des maschinellen Lernens angewendet. Listenbasierten Verfahren liegen Wörterbücher zu Grunde, die vorher - meist in aufwendiger manueller Überarbeitung für das jeweilige Nutzungsszenario - erstellt werden müssen. Für regelbasierte Verfahren werden Muster und Heuristiken erstellt, z.B. weist das Wort "Herr" vor einem Term im Deutschen auf eine Person hin. Diese Regeln müssen zumeist für verschiedene Sprachen unterschiedlich erstellt und bei Wechsel der ursprünglichen Wissensdomäne angepasst werden. Maschinelle Lernverfahren benötigen annotierte Trainingsdaten, um anhand dieser automatisch die Regeln zu extrahieren, die einen bestimmten Datensatz klassifizieren.

Ein Beispiel für eine Anwendung der Eigennamenerkennung ist das Projekt WALU (WIKINGER Annotations- und Lern-Umgebung) [155], das die Entwicklung eines Werkzeugs zum Ziel hat, mit dem Domänenexperten ohne Vorkenntnisse der Computerlinguistik das semantische Auszeichnen von Eigennamen (semi-)automatisch durchführen können. Mit diesem Werkzeug erstellen die Benutzer aus annotierten Textsammlungen Listen von Eigennamen, mit denen maschinelle Lernverfahren trainiert werden, um diese Namen zukünftig automatisch klassifizieren zu können. Neben den maschinellen Lernverfahren werden in WALU auch regelbasierte Methoden angewendet, z.B. reguläre Ausdrücke zur Erkennung von Datums- und Zeitangaben.

Im Gegensatz zur Eigennamenerkennung in natürlichsprachlichen Texten sind Tags simple Zeichenketten. Teilweise können deren Methoden übertragen werden, aber Heuristiken wie "Herr"+Person versagen, wenn das Tag die Anrede nicht enthält, was in den meisten Tagging-Systemen der Fall sein dürfte.

5.2.2 Datenquellen für eine automatische Tag-Typ-Erkennung

In [100] werden die Verfahren der automatischen Erkennung der Tag-Typen eingeteilt in korpusbasierte, wissensbasierte und hybride Ansätze. Korpusbasierte Verfahren nutzen statistische Werte, die auf Basis des zu klassifizierenden Korpus (Tags, Ressourcen etc.) ermittelt werden. Wissensbasierte Ansätze klassifizieren Tags unter Einbezug externer Wissensquellen ähnlich den listenbasierten Verfahren. Statt der manuell erstellten Listen werden bei der Erkennung von Tag-Typen allerdings häufig externe Datenquellen abgefragt.

Der folgende Abschnitt gibt einen Überblick über relevante Datenquellen. Die u.g. Datenquellen sind relevant, da sie bis auf WordNet den Vorteil bieten, dass sie durch große Communities gepflegt werden und weiter wachsen. Themen, die durch aktuelles Weltgeschehen oder Forschung wichtig werden bzw. neu entstehen, werden daher stetig hinzugefügt. Des Weiteren werden in allen Datenquellen die Themen bzw. Objekte in Kategorien eingeordnet. Teilweise verlinken die Einträge in einer Datenbank auf Artikel

in den anderen Datenbanken bzw. auf Kategorien der anderen Datenbanken. Die Datenbanken können frei exportiert werden und stellen verschiedene Möglichkeiten zur Abfrage der Daten zur Verfügung (u.a. Web-Services).

GeoNames

GeoNames¹ ist eine geographische Datenbank, die Informationen zu aktuell mehr als 10 Millionen geographischen Namen enthält, die über 7,5 Millionen topographischen Objekten entsprechen. Die Objekte sind in 9 Klassen (Land/Staat/Region, Fließendes Gewässer/See, Park/Landschaft, Stadt, Straße, Sehenswürdigkeit/Gebäude, Gebirge/Berg, Unterwasser, Wald/Wüste) eingeordnet. Diese 9 Klassen sind in weitere 645 Unterklassen (z.B. Regierungsgebäude, Flughafen, Brücke in der Klasse Sehenswürdigkeit/Gebäude) unterteilt. Die Datenbank enthält die Namen der Orte in der Landessprache sowie deren Übersetzungen in einer Vielzahl anderer Sprachen. Außerdem enthält sie abhängig von der Klasse des Objekts strukturierte Informationen über Einwohnerzahl, Breiten- und Längengrade, Höhe über Meeresspiegel, Postleitzahlen etc. Die Daten werden von GeoNames aus verschiedenen Quellen aggregiert, darunter statistische (Vermessungs-)Ämter verschiedener Länder, Wetterstationen und Wikipedia. Diese Daten sind nicht immer fehlerfrei, können aber durch die Community weiter bearbeitet und bereinigt werden. Als Web-Service stehen beispielsweise Ortssuche über Postleitzahlen oder Umkreissuche zu gegebenen Ortskoordinaten zur Verfügung.

Freebase

Freebase² ist eine große, offene Datenbank, die allgemeines Wissen enthält. Sie umfasst zum jetzigen Zeitpunkt Einträge zu mehr als 20 Millionen Themen bzw. Objekten wie Filmen, Büchern, Personen (2,3 Mio.), Orten (1,8 Mio.), Ereignissen (ca. 100.000) und weiteren. Die Datenbasis wird von der Freebase-Community gepflegt, die Daten aus anderen Quellen aggregiert oder manuell hinzufügt. Für die Einträge existieren strukturierte Informationen wie Geburtsdatum von Personen oder Geokoordinaten von Orten, die durch Schemata beschrieben sind. Teilweise existieren die Artikel in mehreren Sprachen, wobei englische Artikel dominieren auf Grund der Tatsache, dass viele der aggregierten Datenquellen in englischer Sprache verfasst sind. Die Ergebnisse zu einer Suchanfrage werden in Freebase anhand eines Relevanzwertes sortiert, der mittels Ähnlichkeitsbewertung von Anfrage und Artikel in Freebase durch Lucene³ berechnet wird. In die Berechnung fließt die Popularität eines Eintrages (ähnlich PageRank [101]) ein, die sich durch die Links errechnet, die u.a. von anderen Freebase-Artikeln auf diesen Artikel verweisen.

DBpedia

DBpedia⁴ [17] ist ein Community-Projekt, das zum Ziel hat, strukturierte Informationen der Wikipedia zu extrahieren und diese Informationen im Web verfügbar zu machen. Strukturierte Informationen in der

¹ <http://www.geonames.org>, Zuletzt abgerufen am 28.02.2011

² <http://www.freebase.com>, Zuletzt abgerufen am 28.02.2011

³ http://lucene.apache.org/java/2_4_0/api/core/org/apache/lucene/search/Similarity.html, Zuletzt abgerufen am 28.02.2011

⁴ <http://www.dbpedia.org>, Zuletzt abgerufen am 28.02.2011

Wikipedia sind beispielsweise Kategorien und Informationskästen, die abhängig von der Artikelkategorie mit speziellen Templates generiert werden und dadurch größtenteils automatisch auslesbar sind. Informationskästen finden sich z.B. in Artikeln über Städte, die Angaben zu Staat, Einwohnerzahl, Fläche etc. enthalten. Die DBpedia-Datenbank enthält mehr als 3,5 Mio. Einträge (Stand Januar 2010), einschließlich 364.000 Personen, 148.000 Organisationen und 462.000 Orten. Viele Wikipedia-Artikel existieren in mehreren Sprachen, die aufeinander verlinken. Auf Grund der Extraktion dieser Links ist DBpedia multilingual, wenn auch die englische Sprache überwiegt (rund 3,2 Mio. englische Zusammenfassungen, aber nur 0,5 Mio. deutsche). Der Web-Service, der eine Suche in DBpedia zur Verfügung stellt, liefert die passenden Einträge auf Basis des Lucene-Index (siehe Abschnitt "Freebase") gewichtet zurück.

WordNet

WordNet⁵ [72] ist ein Thesaurus in englischer Sprache. Es enthält Substantive, Verben, Adjektive und Adverbien, die in Synonymmengen (Synsets) gruppiert sind, d.h. Gruppen von Wörtern, die untereinander austauschbar sind, ohne den Wahrheitsgehalt der Aussage zu verändern. Diese Synonymmengen sind Kategorien zugeordnet, z.B. Ort, Ereignis oder Nahrungsmittel. Die Synonymmengen enthalten außerdem Verweise innerhalb der Synonymmenge und zu anderen Synonymmengen, z.B. Hyponyme, Antonyme und Holonyme (z.B. Gebäude ist ein Holonym von Fenster). Die aktuelle Version (Stand 2006) enthält 155.287 Wörter und 117.659 Synonymmengen.

Abbildung 5.1 ist ein Screenshot von WordNet zum Suchbegriff "Java". Es wurden 3 Substantive gefunden. Die Insel Java kommt häufiger vor als die anderen beiden Substantive, was durch die Reihenfolge und die Nummern in Klammern dargestellt wird. Die WordNet-Kategorien stehen in spitzen Klammern, z.B. gehört das erste Substantiv zur Kategorie "<noun.location>", ist also demzufolge als Ort klassifiziert. Des Weiteren gibt es eine kurze Beschreibung zu den meisten Wörtern. Das zweite Substantiv bezieht sich auf den Kaffee. Die Synonymmenge für dieses Substantiv ist "coffee, java". Zusätzlich sind Kaffeesorten (z.B. coffee substitute, Irish coffee, cafe au lait,...) als Hyponyme eingeblendet. Bestandteile sind "coffee bean" (Synonymmenge "coffee bean, coffee berry, coffee", Kategorie <noun.food>) und "caffeine" (Synonymmenge "caffeine, caffen", Kategorie <noun.substance>). Das dritte Substantiv bezieht sich auf die Programmiersprache und ist zur Kategorie "<noun.communication>" zugeordnet. Für dieses Substantiv ist sein Hypernym ("object-oriented programming language") und wiederum dessen Hypernym ("programming language") eingeblendet. Jedes dieser Synonyme, Hyperonyme etc. ist ein Link zum eigentlichen Begriff.

Abdeckung der Datenquellen

Die Datenquellen enthalten eine unterschiedliche Anzahl an Konzepten/Termen, so dass es sinnvoll erscheint häufig mehrere Datenquellen für eine automatische Erkennung der Tag-Typen parallel zu verwenden. Beispielsweise werden von WordNet in einer in [139] durchgeführten Analyse nur 52% der untersuchten Flickr-Tags abgedeckt, die restlichen 48% konnten nach Angaben der Autoren nicht in WordNet gefunden werden. WordNet hat nicht das Ziel, aktuelles Weltgeschehen abzubilden; im Gegen-

⁵ <http://wordnet.princeton.edu/>, Zuletzt abgerufen am 28.02.2011

Noun

- (2)<noun.location>[S](#): (n) [Java](#) (an island in Indonesia to the south of Borneo; one of the world's most densely populated regions)
- (1)<noun.food>[S](#): (n) [coffee](#), [java](#) (a beverage consisting of an infusion of ground coffee beans)
 - [direct hyponym](#) / [full hyponym](#)
 - <noun.food>[S](#): (n) [coffee substitute](#) (a drink resembling coffee that is sometimes substituted for it)
 - <noun.food>[S](#): (n) [Irish coffee](#) (sweetened coffee with Irish whiskey and whipped cream)
 - <noun.food>[S](#): (n) [cafe au lait](#) (equal parts of coffee and hot milk)
 - <noun.food>[S](#): (n) [cafe noir](#), [demitasse](#) (small cup of strong black coffee without milk or cream)
 - <noun.food>[S](#): (n) [decaffeinated coffee](#), [decaf](#) (coffee with the caffeine removed)
 - <noun.food>[S](#): (n) [drip coffee](#) (coffee made by passing boiling water through a perforated container packed with finely ground coffee)
 - <noun.food>[S](#): (n) [espresso](#) (strong black coffee brewed by forcing hot water under pressure through finely ground coffee beans)
 - <noun.food>[S](#): (n) [cappuccino](#), [cappuccino coffee](#), [coffee cappuccino](#) (equal parts of espresso and hot milk topped with cinnamon and nutmeg and usually whipped cream)
 - <noun.food>[S](#): (n) [iced coffee](#), [ice coffee](#) (a strong sweetened coffee served over ice with cream)
 - <noun.food>[S](#): (n) [instant coffee](#) (dehydrated coffee that can be made into a drink by adding hot water)
 - <noun.food>[S](#): (n) [mocha](#), [mocha coffee](#) (a superior dark coffee made from beans from Arabia)
 - <noun.food>[S](#): (n) [Turkish coffee](#) (a drink made from pulverized coffee beans; usually sweetened)
 - <noun.food>[S](#): (n) [cafe royale](#), [coffee royal](#) (black coffee with Cognac and lemon peel and sugar)
 - [substance meronym](#)
 - <noun.food>[S](#): (n) [coffee bean](#), [coffee berry](#), [coffee](#) (a seed of the coffee tree; ground to make coffee)
 - <noun.substance>[S](#): (n) [caffeine](#), [cafein](#) (a bitter alkaloid found in coffee and tea that is responsible for their stimulating effects)
 - [direct hypernym](#) / [inherited hypernym](#) / [sister term](#)
- <noun.communication>[S](#): (n) [Java](#) (a platform-independent object-oriented programming language)
 - [direct hypernym](#) / [inherited hypernym](#) / [sister term](#)
 - <noun.communication>[S](#): (n) [object-oriented programming language](#), [object-oriented programing language](#) ((computer science) a programming language that enables the programmer to associate a set of procedures with each type of data structure)
 - <noun.communication>[S](#): (n) [programming language](#), [programing language](#) ((computer science) a language designed for programming computers)

[WordNet home page](#)

Abbildung 5.1: Screenshot von WordNet zum Suchbegriff "Java"

satz zu Freebase und DBpedia enthält es, wie die Zahlen oben zeigen, weit weniger Terme. Des Weiteren hat es den Stand von 2006 und enthält daher viele in der Zwischenzeit neu entstandene Themen nicht.

5.2.3 Maße für den Vergleich von Klassifikationsverfahren

Die Aufgabe von Klassifikationsverfahren ist die Zuordnung von Objekten zu vorgegebenen Kategorien. Die Bestimmung des Tag-Typs, d.h. die Zuordnung eines Tags zu einem der definierten Tag-Typen, entspricht daher einer Klassifikationsaufgabe. Ein Klassifikationsverfahren macht bei der Zuordnung in der Regel Fehler, d.h. es ordnet manche Objekte einer falschen Klasse zu. Im Falle der Erkennung des Tag-Typs ist es ein Fehler, wenn z.B. eine Person als Ort klassifiziert wird. Für den Vergleich von Klassifikationsverfahren gibt es verschiedene Maße. Für die Evaluation eines Verfahrens, d.h. die Bestimmung der Maße, müssen die korrekten Kategorien der Objekte bekannt sein. Dadurch können die Fehler anhand der vom Verfahren vorgenommenen Einordnung im Vergleich zur korrekten Einordnung ermittelt

werden. Dabei werden zwei Fehler unterschieden: "False Positives" und "False Negatives". False Positives sind Objekte, die fälschlicherweise zu dieser Klasse zugeordnet werden und False Negatives sind Objekte, die trotz ihrer tatsächlichen Zugehörigkeit zur Klasse nicht in diese Klasse eingeordnet wurden. Zum Vergleich von Verfahren für die Klassifizierung werden aus der relativen Häufigkeit dieser Fehler oftmals die folgenden weit verbreiteten Maße berechnet [11, 88]: Precision, Recall, F_1 -Measure (harmonisches Mittel von Genauigkeit und Trefferquote) und Accuracy.

Der Recall gibt das Verhältnis aus korrekt klassifizierten Objekten einer Kategorie (true Positives) und der Gesamtanzahl der relevanten Objekte dieser Kategorie an und berechnet sich wie folgt:

$$Recall = \frac{TruePositives}{TruePositives + FalseNegatives}$$

Der Wertebereich des Recalls erstreckt sich von 0 bis 1, wobei 0 das schlechteste Ergebnis und 1 das bestmögliche Ergebnis ist, d.h. 1 wird erreicht, wenn alle relevanten Objekte korrekt klassifiziert wurden.

Die Precision gibt das Verhältnis aus korrekt klassifizierten Objekten einer Kategorie und der Gesamtanzahl aller in die Kategorie eingeordneten Objekte an und berechnet sich wie folgt:

$$Precision = \frac{TruePositives}{TruePositives + FalsePositives}$$

Der Wertebereich der Precision erstreckt sich ebenfalls von 0 bis 1, wobei auch hier 1 das bestmögliche Ergebnis darstellt.

Damit eine Aussage über die Güte des Systems getroffen werden kann, sollten immer beide Maße berechnet werden. Denn ein Recall von 1 in einer Kategorie kann bereits dadurch erreicht werden, dass alle Objekte in diese Klasse eingeordnet werden und damit alle relevanten Objekte zwar gefunden werden, aber auch sehr viele falsch klassifizierte Objekte zu dieser Kategorie zugeordnet werden (was damit zu geringen Precision-Werten führt).

Ein typisches Maß, das sowohl Precision als auch Recall einbezieht, ist das F_1 -Measure. Beim F_1 -Measure werden Precision und Recall gleich gewichtet. Häufig wird beim F_1 -Measure der Index 1 weggelassen. In Anwendungsfällen, in denen entweder ein hoher Recall oder eine hohe Precision wichtig ist, können die Maße auch unterschiedlich gewichtet werden. Das F_1 -Measure berechnet sich wie folgt:

$$F_1 = 2 * \frac{precision * recall}{precision + recall}$$

Die Accuracy gibt das Verhältnis aus allen richtig klassifizierten und allen evaluierten Objekten an. Es berechnet sich wie folgt:

$$Accuracy = \frac{TruePositives + TrueNegatives}{TruePositives + FalsePositives + FalseNegatives + TrueNegatives}$$

Klassifikationsaufgaben werden häufig mittels maschinellen Lernverfahren gelöst. Für maschinelle Lernverfahren ist es üblich, die vorhandene Menge von Objekten in zwei Mengen aufzuteilen. Eine dieser Mengen dient als Trainingsmenge (Trainingskorpus), mit der der Algorithmus trainiert wird und die andere Menge (Testkorpus) dient zur Messung der Qualität des Algorithmus. Die Aufteilung der Objekte in diese beiden Mengen hat Einfluss auf die Ergebnisse des Algorithmus, daher wird die Aufteilung in Trainings- und Testmenge mehrfach durchgeführt und der Durchschnitt über alle Ergebnisse ermittelt. Dies wird als n-fache Kreuzvalidierung bezeichnet, bei einer 10-fachen Kreuzvalidierung wird beispielsweise der Korpus in 10 Teile aufgeteilt und jeder dieser Teile dient jeweils einmal als Trainingskorpus.

Bei der Zusammenstellung einer Menge von Objekten zur Evaluierung eines Klassifikationsverfahrens muss die Zuordnung von Objekten zu Kategorien in der Regel manuell vorgenommen werden. In einigen Fällen kann die Zuordnung subjektiv sein, d.h. von der Person, die die Zuordnung vornimmt, abhängen. In solchen Fällen sollte die Zuordnung von verschiedenen Personen durchgeführt werden. Für die Beurteilung der Übereinstimmung von Personen werden Maße wie Cohen's Kappa [33] oder Fleiss' Kappa_{free} [113] verwendet, womit die sogenannte Interrater-Reliabilität berechnet werden kann. Anhand von Wertebereichen, wie in Tabelle 5.1 gezeigt, kann die Verlässlichkeit der Zuordnungen eingeschätzt werden. Die Berechnung des Fleiss' Kappa_{free} wird empfohlen, wenn mehr als zwei Personen manuell klassifizieren und vorher nicht wissen, wie groß die einzelnen Klassen sind [20].

Kappa κ	Grad der Übereinstimmung
$\kappa < 0$	Schlechte Übereinstimmung
$0 < \kappa < 0,20$	Etwas Übereinstimmung
$0,21 < \kappa < 0,40$	Ausreichende Übereinstimmung
$0,41 < \kappa < 0,60$	Mittelmäßige Übereinstimmung
$0,61 < \kappa < 0,80$	Beachtliche Übereinstimmung
$0,81 < \kappa < 1,00$	(Fast) vollkommene Übereinstimmung

Tabelle 5.1: Beurteilung der Interrater-Reliabilität nach [80]

5.2.4 Existierende Ansätze zur Klassifikation von Tag-Typen

Es existieren bisher wenige Arbeiten im Bereich der automatischen Klassifizierung von Tag-Typen. Diese Arbeiten verfolgen das Ziel der Klassifizierung von Tag-Typen jeweils bezogen auf Bilder, Webseiten und/oder Musik. Ein Überblick über bis 2009 existierende Ansätze wird in [16] gegeben. Einige der Ansätze nutzen nicht nur Tags, sondern beziehen zusätzliche Metadaten wie z.B. das Aufnahmedatum der Bilder oder den Inhalt einer getaggten Webseite in den Klassifikationsalgorithmus ein. Die verwandten, seit 2009 veröffentlichten Arbeiten sind in Tabelle 4.1 auf Seite 53 mit ihren klassifizierten Tag-Typen aufgelistet. Die Arbeit von Bischoff et al. analysiert Bilder, Webseiten und Musik, der Ansatz von Warrena [156] Bücher, und das Verfahren in TagExplorer [140] Bilder. Die Tag-Typen der einzelnen Ansätze sind nicht identisch, aber überlappen sich, wie die Analyse der Tag-Typen in Abschnitt 4.3.1 zeigt.

Zur Klassifikation der Webseiten werden alle drei in 5.2.1 genannten Methoden der Eigennamenklassifikation verwendet. Für den Typ Zeit werden reguläre Ausdrücke und von den Autoren manuell erstellte Listen genutzt, die z.B. Feiertage, Wochentage, Jahreszeiten usw. enthalten. Eine ebenfalls manuell erstellte Liste enthält 83 bekannte Dateiformate für Typ/Medium. Für Selbstreferenzen wurde eine Liste initial mit 28 Stichwörtern erstellt, die für die Delicious-Daten noch um Wörter wie "homepage" und "login" ergänzt wurde. Eine Liste mit Orten wurde aus dem geographischen Thesaurus GATE⁶ extrahiert, um den Tag-Typ Ort zu identifizieren. Diese Orts-Liste wurde manuell um ca. 120 Orte (z.B. "java", "nice", "church") bereinigt, da diese Orte auch sehr allgemeine Wörter darstellen und vom Verfahren daher nicht als Ort erkannt werden sollen. Ein Tag wird in dem Verfahren als Autor klassifiziert, wenn der Bezeichner des Tags in der URL der Webseite enthalten ist. Diese fünf Typen werden mit den erstellten Listen oder regulären Ausdrücken abgeglichen. Tags, die keinem dieser fünf Tag-Typen zugeordnet werden können, werden mit Hilfe von maschinellen Lernverfahren den anderen drei Tag-Typen zugeordnet. Für das maschinelle Lernen werden dabei folgende Features genutzt: Anzahl der Nutzer bzw. Tag-Häufigkeit, Anzahl der Wörter, Anzahl der Zeichen, Wortart und WordNet-Kategorien. Die Autoren treffen die Annahmen, dass Themen überwiegend aus Substantiven bestehen und Meinungen überwiegend aus Adjektiven und einer höheren Anzahl an Wörtern. Die Häufigkeit für Tags, die zum Tag-Typ Nutzungskontext gehören, bezogen auf die Menge aller Tags, wird als sehr gering eingeschätzt. Bis auf die URL nutzt dieser Ansatz keine zusätzlichen Metadaten oder Inhalte der getaggten Webseiten. Falls für ein Tag mehrere Bedeutungen in WordNet existieren, wird versucht Synonyme anhand von co-occurring Tags zu bestimmen. Co-occurring Tags sind Tags, die gemeinsam mit diesem Tag vorkommen. Falls auch die Synonyme keine eindeutige Bedeutung in WordNet haben, wird die in WordNet gängigste Bedeutung und deren Kategorien verwendet.

Für die Evaluation wurde eine Menge von 700 Delicious-Tags zusammengestellt, darunter die 300 häufigsten Tags. Diese Tags und weitere 1400 Tags für Bilder und Musik wurden manuell klassifiziert. Die Ergebnisse des Verfahrens der Delicious-Tags werden in Tabelle 5.2 dargestellt. Die Tabelle listet neben Precision und Recall auch die verwendeten Features für jede Klasse und die prozentuale Verteilung der Tag-Typen auf. Beispielsweise werden von diesem Ansatz für den Tag-Typ Thema ein Recall von 96,38% und eine Precision von 83,89% erreicht. Meinungen und Tags zum Nutzungskontext in der Menge der untersuchten Tags werden von diesem Ansatz nicht erkannt. 11,57% der untersuchten Tags können vom System nicht klassifiziert werden. Insgesamt ergibt sich über alle Kategorien anhand der Verteilung der Tags gewichtet ein F-Measure von 69,15%.

Wartena

Wartena versucht die Tags mit Verfahren des maschinellen Lernens zu klassifizieren und nutzt dafür verschiedene Features, die er vom Korpus (Tags, Ressourcen, Nutzer etc.) ableitet. Dies sind: Exzentrizität (Eccentricity) bezogen auf Korpus oder Nutzer, Tag-Häufigkeit, Bewertungen des Buches durch die Nutzer, Namen der Autoren des Buches und das Enthaltensein von bestimmten Zeichenketten (dies sind "edition", "author", "reading", "read", "great", "prize", "award" und "favorite"). Exzentrizität ist ein Maß

⁶ <http://gate.ac.uk>, Zuletzt abgerufen am 28.02.2011

Tag-Typ	Features	Precision [%]	Recall [%]	Häufigkeit [%]	F-Measure [%]
Thema	Wortart, WordNet-Kategorien	83,89	96,38	67,14	89,70
Zeit	Reguläre Ausdrücke, Listen	100,00	100,00	0,86	100,00
Ort	Listen	70,37	70,37	3,86	70,37
Typ	Listen	66,67	42,86	8,00	52,18
Autor	Reguläre Ausdrücke, Listen	9,85	38,57	6,29	15,69
Meinung	Anzahl Wör- ter, Anzahl Zeichen, Wortart, WordNet-Kategorien	0,00	0,00	5,14	0,00
Nutzungs- kontext	Wortart, WordNet-Kategorien	0,00	0,00	7,86	0,00
Selbstreferenz	Listen	33,33	16,67	0,86	22,22

Tabelle 5.2: Ergebnisse der Klassifizierung von Delicious-Tags vgl. [16]

für die Abweichung der Co-Occurrence-Verteilung von der allgemeinen Tag-Verteilung; eine hohe Exzentrizität ergibt sich für Tags, die mit wenigen verschiedenen Tags gleichzeitig für Ressourcen vergeben werden.

Klasse (2)	Klasse (5)	Anzahl
Werk	Thema	150
	Autor	100
	Attribut	150
Nutzer	Selbstreferenz	100
	Meinung	65

Tabelle 5.3: Anzahl der ausgewählten Tags je Klasse [156]

Für die Evaluation des Verfahrens wurden 565 Tags zufällig ausgewählt und manuell den vorgegebenen Klassen zugeordnet. Die Verteilung der Tags auf die Klassen ist in Tabelle 5.3 aufgelistet. Tabelle 5.4 enthält die Ergebnisse der 10-fachen Kreuzvalidierung, jeweils für die Zuordnung zu 5 und zu 2 Klassen (Klassifikation in "Werk" und "Nutzer"). Die Baseline entspricht hier einem Klassifikations-Algorithmus, der alle Tags einfach zur Klasse mit der höchsten Wahrscheinlichkeit (d.h. in diese Klasse gehören die meisten Objekte) zuordnen würde. In diesem Fall hat die Klasse "Thema" die meisten Objekte und zwar 27% aller zu klassifizierenden Objekte, demzufolge beträgt die abgeleitete Baseline 27%. Für die Klassi-

fikation in nur zwei Klassen werden die Unterklassen zusammengefasst und es ergibt sich eine Baseline von 71%.

Bei Verwendung aller Features werden 75% bzw. 92% der Tags richtig klassifiziert. Allerdings gibt der Autor an, dass das Feature für das Enthaltensein bestimmter Zeichenketten großen Einfluss auf das Ergebnis hat, sodass die Anzahl prozentual richtig klassifizierter Tags stark schwankt. Bei Weglassen dieses Features werden noch 67% bzw. 89% der Tags richtig klassifiziert.

	5 Klassen	2 Klassen
Baseline	27%	71%
Alle Features	75%	92%
Ohne das Feature "Enthalten-sein von bestimmten Zeichenketten"	67%	89%

Tabelle 5.4: Ergebnisse der prozentual richtig klassifizierten Tags [156]

TagExplorer

Im TagExplorer werden die Tag-Typen mittels WordNet identifiziert. Dazu sind die 5 Tag-Typen des Tag-Explorers auf bestimmte WordNet-Kategorien gemappt, siehe Abbildung 5.2. Beispielsweise werden die WordNet-Kategorien "<noun.person>" und "<noun.group>" auf den Tag-Typ Person/Gruppe gemappt. Der zu klassifizierende Tag-Bezeichner wird in WordNet gesucht und anhand der WordNet-Kategorie des ersten Substantivs wird der Tag-Typ abgeleitet. Für das Beispiel "Java" würde WordNet die Substantive entsprechend Abbildung 5.1 auf Seite 67 zurückliefern. Das Verfahren im TagExplorer würde das erste Substantiv, d.h. die Insel auswählen und dessen WordNet-Kategorie ist "Ort". Deshalb würde der Tag-Explorer "Java" als Ort klassifizieren. Die Autoren treffen die Annahmen, dass die WordNet-Kategorien korrekt sind, und dass das häufigste Substantiv in WordNet dem Tag am wahrscheinlichsten entspricht.

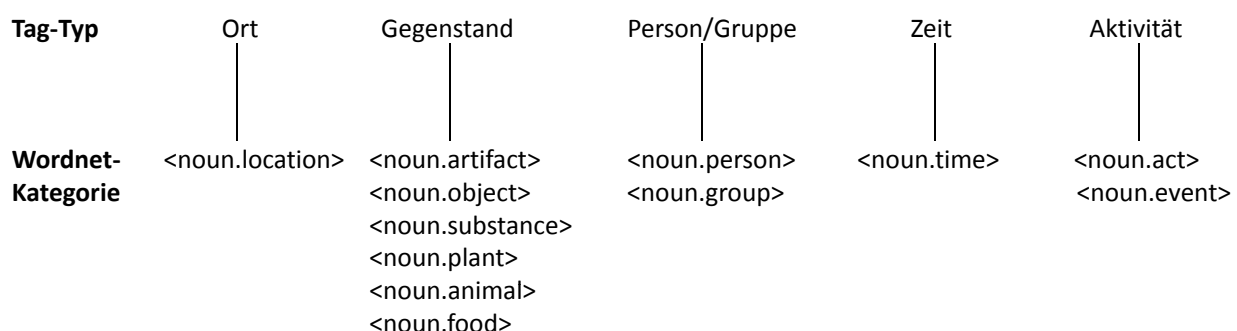


Abbildung 5.2: Mapping von Tag-Typen und WordNet-Kategorien im TagExplorer (vgl. [140])

Da mit einem direkten Vergleich von Tags und WordNet-Begriffen nach Angaben der Autoren nur 52% der untersuchten Flickr-Tags zu einer WordNet-Kategorie zugeordnet werden können, nutzt der

TagExplorer auch das ClassTag-System [100]. Das Vorgehen des ClassTag-Systems ist in Abbildung 5.3 dargestellt.

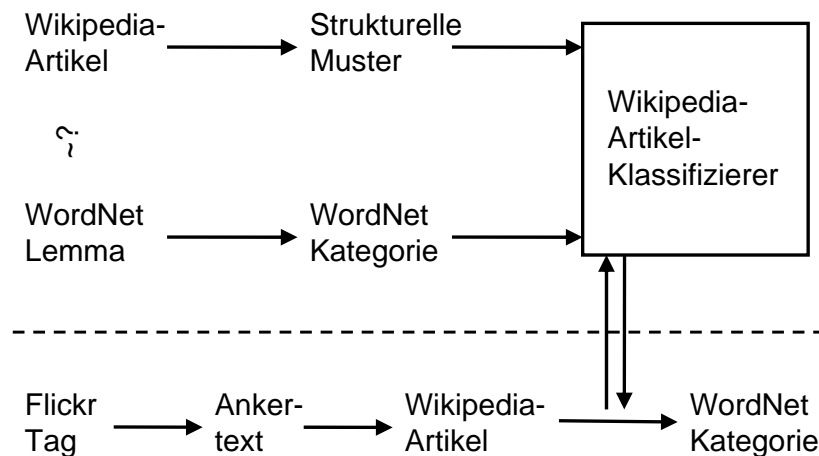


Abbildung 5.3: Vorgehen im ClassTag-System (vgl. [100])

Mittels Wikipedia versucht das ClassTag-System die WordNet-Kategorien weiterer Tags zu identifizieren. Für diese Zuordnung sind drei Schritte notwendig (siehe unterer Teil der Abbildung). Der erste Schritt (Flickr-Tag) versucht durch Zeichenkettenvergleich einen Ankertext zu finden. Ankertexte entsprechen Texten von Links in Wikipedia-Artikeln. Der zweite Schritt ermittelt die passenden Wikipedia-Artikel zum Ankertext. Der dritte Schritt identifiziert mittels maschinellem Lernen anhand der Wikipedia-Kategorien und Wikipedia-Templates (z.B. Informationskästen) die passende WordNet-Kategorie. Für diesen dritten Schritt (siehe oberer Teil der Abbildung 5.3) gibt es verschiedene Parameter, die Recall und Precision des ClassTag-Systems und demzufolge auch den TagExplorer beeinflussen. Mit einer Precision von 72% können beispielsweise für 39% der analysierten Wikipedia-Artikel eine korrekte WordNet-Kategorie zugeordnet werden.

Am Beispiel der Tags "chryslerbuilding" und "nyc" soll das Vorgehen beispielhaft in Abbildung 5.4 erläutert werden. Für das Tag "chryslerbuilding" konnte der Ankertext "Chrysler Building" identifiziert werden. Dieser Ankertext verweist 188-mal auf den Wikipedia-Artikel "Chrysler Building". Das maschinelle Lernverfahren klassifiziert den Artikel mit einem Gewicht von 0.83 als WordNet-Kategorie "Artefakt". Für "nyc" wurden 2 Ankertexte gefunden. Der erste verweist 40-mal auf den Wikipedia-Artikel "New York City" und der zweite verweist 85-mal auf diesen Wikipedia-Artikel und 7-mal auf den Wikipedia-Artikel "New York". Für beide Wikipedia-Artikel ergibt das Verfahren die WordNet-Kategorie "Location", d.h. dieses Tag würde anschließend im Mapping der WordNet-Kategorien auf TagExplorer-Kategorien als "Ort" klassifiziert werden.

Laut einer Hochrechnung der Autoren lassen sich mit diesem Verfahren 69% aller Flickr-Tags klassifizieren. Die Precision dieser Klassifikation ist von den Autoren nicht direkt angegeben. Für Tags, die direkt in WordNet gefunden werden, nehmen die Autoren eine korrekte Klassifikation an. Für die Tags, die nur über "den Umweg in Wikipedia" klassifiziert werden, entspricht die Precision dem ClassTag-System, d.h. 72%. Die Autoren vermuten, dass ein großer Teil der unklassifizierten Tags nicht-englischsprachige Tags

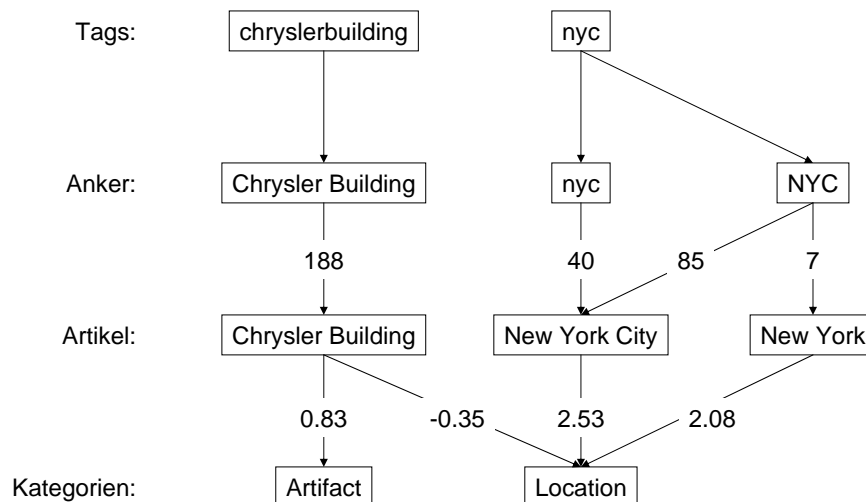


Abbildung 5.4: Vorgehen im ClassTag-System: Tag → WordNet-Kategorie (vgl.[100])

sind, die sie mit ihrem Ansatz nicht abdecken können. In Abbildung 5.5 ist die Verteilung der Flickr-Tags für die verschiedenen Tag-Typen illustriert. Tag-Disambiguation, um beispielsweise zwischen "Java" als Insel und als Kaffee zu unterscheiden, wird bisher nicht durchgeführt, da Maße wie Co-Occurences von Tags bisher nicht im Verfahren genutzt werden.

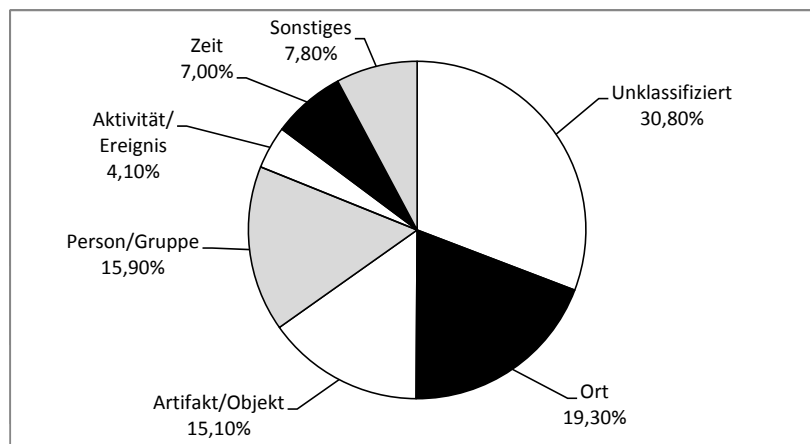


Abbildung 5.5: Verteilung der Tag-Typen in Flickr (vgl. [140])

5.3 Ein wissensbasiertes Verfahren zur Erkennung von Tag-Typen

Im Folgenden soll das im Rahmen dieser Arbeit entwickelte Verfahren zur Erkennung der Tag-Typen und dessen Evaluationsergebnisse für verschiedenen Korpora beschrieben werden. Im Gegensatz zu den vorgestellten verwandten Ansätzen basiert das Verfahren hauptsächlich auf der Abfrage von externen Wissensquellen statt eines korpusbasierten Ansatzes.

5.3.1 Verwendete Datenquellen und Vorgehen

Die Tag-Typen in dieser Arbeit überlappen sich teilweise mit den Tag-Typen der verwandten Ansätze, jedoch gibt es in dem Anwendungsszenario des Ressourcen-basierten Lernens einige Besonderheiten. Der Tag-Typ Ziel ist von der automatischen Klassifikation ausgeschlossen, da die Benutzer durch das Zielmanagement angeregt werden, Ziele vor der Recherche anzulegen, so dass beim Tagging nur bereits existierende Zieltags verwendet werden. Meinungen und Selbstreferenzen werden nicht klassifiziert, da diese als nicht wichtige Tag-Typen identifiziert wurden (siehe Abschnitt 4.3.1) und daher in dieser Arbeit entwickelten Unterstützungswerkzeug ELWMS.KOM nicht enthalten sind.

Bei der Erkennung des Tag-Typs gibt es zwei grundlegende Herausforderungen: Zum einen muss zunächst herausgefunden werden, welche Tag-Typen für das zu klassifizierende Tag gültig sind (z.B. kann das Tag "Merkel" sich sowohl auf eine Person als auch einen Ort in Texas beziehen und damit wäre die Zuordnung zu mindestens zwei Tag-Typen valide). Dies wird im Ansatz durch die Nutzung typenspezifischer Datenbanken (GeoNames, extrahierte Listen und eine Personendatenbank [86]) abgedeckt. Zum anderen muss das Verfahren entscheiden, welcher Tag-Typ für das zu klassifizierende Tag am wahrscheinlichsten ist, d.h. welches Konzept mit dem Bezeichner des Tags mit hoher Wahrscheinlichkeit gemeint ist. Hierzu wird die externe allgemeine Datenquelle Freebase genutzt. Die Tag-Typen, die von diesem Ansatz erkannt werden sollen, lassen sich auf Freebase-Kategorien abbilden, die den Einträgen in Freebase zugeordnet sind. So lassen sich Freebase-Suchergebnisse auf die verschiedenen Tag-Typen abbilden. Dem Ansatz liegt die Annahme zu Grunde, dass Freebase-Einträge, die im Suchergebnis für die Suche nach einem Tag weiter vorn stehen eine höhere Plausibilität besitzen, und der mittels Kategorienabbildung ermittelte Tag-Typ wahrscheinlich gemeint ist.

Weitere im Verfahren genutzte externe Datenquellen sind DBpedia, Wikipedia, WordNet, die Wortliste des deutschen Wörterbuches Beolingus⁷ und Stoppwort-Listen⁸.

In Abbildung 5.6 ist das Vorgehen zur Identifikation des Tag-Typs schematisch dargestellt. Eingabeparameter ist der vom Nutzer eingegebene Tag-Bezeichner. Zuerst findet eine Vorverarbeitung (Schritt 1) des Tag-Bezeichners statt. Anschließend wird Freebase abgefragt (Schritt 2) und dessen Ergebnis an die verschiedenen Klassifizierer (Schritt 3-7) übergeben. Diese gewichten daraufhin parallel unter Zuhilfenahme weiterer typenspezifischer Datenquellen und auf Basis von in dieser Arbeit entwickelten Heuristiken ihren jeweiligen Tag-Typ, falls sie den Tag-Typ zuordnen können. Anschließend werden die Konzepte der verschiedenen Klassifizierer anhand ihres Ranges sortiert (Schritt 8). Ausgabeparameter des Verfahrens ist dann der Tag-Typ, der für den Tag-Bezeichner am wahrscheinlichsten in Frage kommt, oder auch eine gerankte Liste von Tag-Typen. Im Folgenden wird jeder Schritt des Verfahrens erläutert.

Vorverarbeiten (1)

Dieser Schritt dient dazu, den Tag-Bezeichner in eine Form zu bringen, die als Suchbegriff für die externen Quellen dienen kann. In manchen Tagging-Systemen wie Delicious entspricht das Leerzeichen dem Trennzeichen für Tags, d.h. Schlagworte können beim Tagging hintereinander eingegeben werden

⁷ <http://dict.tu-chemnitz.de/>, Zuletzt abgerufen am 28.02.2011

⁸ <ftp://ftp.cs.cornell.edu/pub/smart/english.stop>, Zuletzt abgerufen am 28.02.2011 für englischsprachige Stoppwörter und <http://solariz.de/wp-content/files/stopwords.txt>, Zuletzt abgerufen am 28.02.2011 für deutschsprachige Stoppwörter

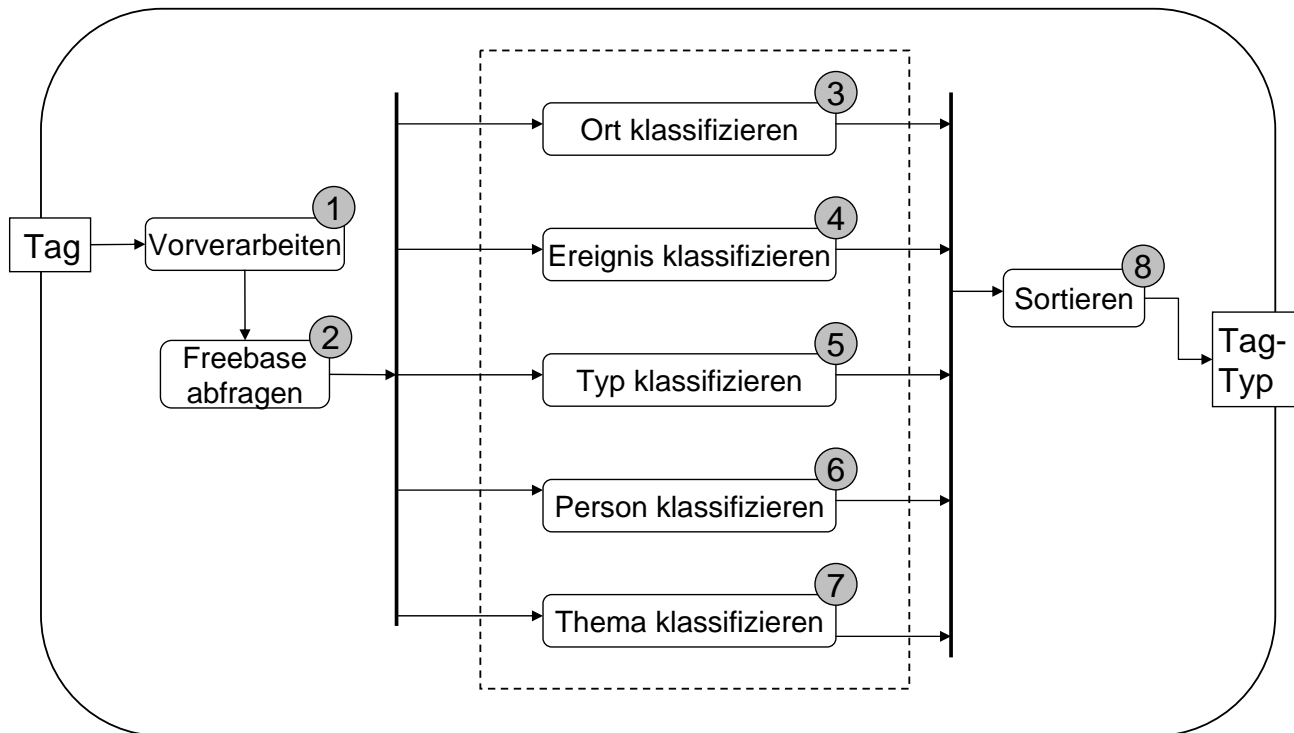


Abbildung 5.6: Verfahren zur automatischen Erkennung der Tag-Typen

und das Leerzeichen trennt diese dann in einzelne Tag-Bezeichner. Dies macht die Verwendung von mehreren Termen in einem Tag-Bezeichner unmöglich. Daher werden von manchen Nutzern Zeichen wie "_" oder "." verwendet, um diese Trennung zu umgehen und ein Leerzeichen zu "simulieren". Diese Sonderzeichen werden im Schritt der Vorverarbeitung durch Leerzeichen ersetzt. Eine andere, verwendete Möglichkeit zur Simulation der Leerzeichen ist die CamelCase-Schreibweise. Der Anfangsbuchstabe jedes Wortes wird in dieser Schreibweise groß geschrieben, um den Beginn eines neuen Wortes sichtbar zu machen. Kommen im Tag-Bezeichner Großbuchstaben vor, werden daher vor diesem Großbuchstaben jeweils Leerzeichen eingefügt. Des Weiteren wird eine Stoppwortfilterung durchgeführt.

Da einige der Tags in deutscher Sprache vorkommen, aber die verwendeten Datenquellen teilweise nur englischsprachige Konzepte enthalten, werden diese Tags mittels Wikipedia übersetzt. Die Artikel der deutschen Wikipedia enthalten oft einen Link zur englischen Wikipedia. Existiert der Tag-Bezeichner als Artikel in der Wikipedia, wird daher das verlinkte Konzept der englischen Wikipedia als Übersetzung verwendet.

Freebase abfragen (2)

Der vorverarbeitete Tag-Bezeichner und falls vorhanden auch dessen Übersetzung werden als Suchbegriffe genutzt, um Freebase abzufragen. Die Abfrage von Freebase liefert eine sortierte Menge von Konzepten zurück, die verschiedenen Freebase-Kategorien zugeordnet sind. Die Sortierung basiert, wie in Abschnitt 5.2.2 beschrieben, auf der Häufigkeit der Verwendung des Konzeptes. Im Rahmen des in dieser Arbeit entwickelten Klassifikations-Algorithmus wird die Annahme getroffen, dass die Position in der Freebase-Liste über die Plausibilität eines Tag-Typs entscheidet, d.h. je weiter vorn das Konzept in

der Trefferliste eingeordnet ist, desto wahrscheinlicher ist es, dass sich der Tag-Bezeichner auf dieses Konzept bezieht. "Merkel" als Freebase-Konzept für die Bundeskanzlerin Angela Merkel wird beispielsweise als erster Treffer zurückgeliefert, während der Ort Merkel in Texas erst an dritter Stelle auftaucht. Diese Position in der Trefferliste von Freebase wird als Rang bei jedem Eintrag in der Liste gespeichert.

Anhand der Freebase-Kategorien, die auf die Tag-Typen abgebildet werden, wird die Treffermenge für jeden Klassifizierer unterteilt, d.h. der Orts-Klassifizierer z.B. erhält die Menge aller gefundenen Orte. Einige der Freebase-Kategorien können nicht auf die Tag-Typen abgebildet werden, dessen Konzepte werden an keinen Klassifizierer weitergegeben, sondern verworfen. Ausgeschlossen werden beispielsweise Konzepte, die einem Liedtitel oder einem Buchtitel entsprechen.

Jeder Eintrag auf der Freebase-Liste wird auf vollständige Identität zwischen Tag-Bezeichner und Freebase-Bezeichner geprüft. Des Weiteren wird bei Freebase-Konzepten, die dem Tag-Typ Person entsprechen (siehe Schritt 6) geprüft, ob der Tag-Bezeichner in dem Freebase-Bezeichner enthalten ist. Beim Vergleich der Bezeichner werden die Großbuchstaben in Kleinbuchstaben umgewandelt und alle Sonderzeichen entfernt. Freebase findet Einträge auch anhand ihrer Abkürzung. Wird der Tag-Bezeichner als Abkürzung des gefundenen Eintrages identifiziert z.B. "nyc" als Abkürzung von "New York City", wird zu diesem Eintrag in der Liste gespeichert, dass es sich um eine Abkürzung handelt.

Tag-Bezeichner, mögliche Übersetzung und die Freebase-Liste der Konzepte, die auch den Rang enthält, werden an die einzelnen Klassifizierer übergeben.

Ort klassifizieren (3)

Der Orts-Klassifizierer fragt zusätzlich zu Freebase die externe Datenquelle GeoNames ab, die sehr viele Orte und deren Übersetzungen in viele Sprachen enthält. Das Vorhandensein des Tag-Bezeichners sowohl in Freebase als auch in GeoNames weist daraufhin, dass das Tag ein Ort ist. Dabei ist zu beachten, dass der Orts-Bezeichner identisch zum Tag-Bezeichner sein muss. Der Orts-Klassifizierer gibt dann diesen Ort mit dem Rang von Freebase zurück. Beispielsweise wird der Ort "Merkel" in beiden Quellen als Ort gefunden. Die Position in Freebase ist 3, so dass ein Rang von 3 für den Tag-Typ Ort gespeichert wird.

Wird der Ort nicht in Freebase, aber in GeoNames gefunden, handelt es sich möglicherweise um eine nicht-englischsprachige Bezeichnung des Ortes. Da es in GeoNames für viele Orte eine Übersetzung gibt, wird versucht, die englische Übersetzung des Ortes in GeoNames zu extrahieren, um mit dieser Freebase erneut abzufragen. Falls mit der Übersetzung dann auch in Freebase ein Ort gefunden werden kann, wird dieser mit seinem Freebase-Rang vom Orts-Klassifizierer zurückgegeben. Bei dem Tag "deutschland" handelt es sich z.B. um eine deutschsprachige Bezeichnung. In Freebase wird dies zwar als Ort gefunden, aber mit der Bezeichnung "Bundesrepublik Deutschland", so dass es von der Freebase-Liste entfernt wird. Die englische Übersetzung in GeoNames lautet "Germany", was in Freebase als Ort gefunden werden kann.

Ereignis klassifizieren (4)

Aus der Tabelle der Feiertage in Deutschland in Wikipedia⁹ wurden deutschsprachige Feiertage (z.B. Christi Himmelfahrt) extrahiert. Entspricht der Tag-Bezeichner einem dieser Feiertage, wird er als Ereignis vom Ereignis-Klassifizierer mit dem Rang 0 zurückgegeben.

⁹ http://de.wikipedia.org/wiki/Feiertage_in_Deutschland, Zuletzt abgerufen am 28.02.2011

Enthält der Tag-Bezeichner eine vierstellige Zahl, wird er als Ereignis mit dem Rang 0 klassifiziert, z.B. "hypertext 2010".

Entspricht der Tag-Bezeichner einer Kombination aus einer Zahl zwischen 1 und 31 und einem Monatsnamen, wird er als Ereignis mit dem Rang 0 klassifiziert, z.B. 4. April.

Trifft keines dieser Kriterien zu, wird die Liste möglicher Freebase-Ereignisse untersucht.

Typ klassifizieren (5)

Im Web existiert eine Liste¹⁰ von Webgenres, die von Autoren im Bereich der Webgenre-Identifikation zusammengestellt wurde. Diese Webgenres wurden extrahiert und mittels Wikipedia automatisch übersetzt (z.B. Commentary in Kommentar). Wenn der Tag-Bezeichner in dieser Liste der (übersetzten) Webgenres enthalten ist, wird er als Typ mit dem Rang 0 klassifiziert. Ist er nicht enthalten, wird Freebase abgefragt.

Person klassifizieren (6)

Folgt der Tag-Bezeichner dem Muster "[Buchstabe + Wort]" oder "[Wort + Buchstabe]" wird er als Person mit dem Rang 0 klassifiziert, z.B. "C Rensing".

Freebase wird analog zu den anderen Klassifizierern abgefragt, mit der Ausnahme, dass keine Übersetzung verwendet wird, da Personen bzw. Organisationen in der Regel keine Übersetzung haben. Bei der Prüfung der Freebase-Einträge wird nicht nur auf vollständige Identität von Tag-Bezeichner und Personennamen geprüft, sondern auch ob der Tag-Bezeichner in dem Personennamen enthalten ist, z.B. "rensing" ist teilweise in "Christoph Rensing" enthalten. Denn Tags bestehen häufig nur aus dem Nachnamen einer Person, z.B. Rensing.

Freebase enthält nur bekannte Personen der Zeitgeschichte, beispielsweise Politiker oder Sportler. Personen, die eine gewisse Berühmtheit nicht überschritten haben, können in dieser Quelle nicht gefunden werden. Daher wird für die Klassifizierung dieser Personen noch eine Datenbank abgefragt, die Vornamen und Familiennamen in Deutschland bzw. USA enthält. Je häufiger der Tag-Bezeichner als Vor- oder Nachname existiert, desto höher ist die Wahrscheinlichkeit, dass das Tag einer Person entspricht, z.B. bei "Anna Meier" oder "Karl Müller", und desto niedriger wird der Rang für den Tag-Typ gesetzt.

Thema klassifizieren (7)

Der Thema-Klassifizierer fragt Freebase analog zu den anderen Klassifizierern ab. Alle Freebase-Konzepte, die keinem der anderen Tag-Typen zugehören und keinem Ausschlusskriterium (z.B. Liedtitel) entsprechen, werden dem Tag-Typ Thema zugeordnet.

Wird der Tag-Bezeichner nicht in Freebase gefunden, existiert aber als Titel einer deutschen Wikipedia-Seite, wird dieses Tag aufgrund der nutzerübergreifenden Bedeutung als Thema mit dem Rang 101 klassifiziert. Dieser Rang ist höher als der maximale Freebase-Rang, denn es werden maximal die 100 ersten Freebase-Treffer berücksichtigt. Diese Heuristik trifft somit nur zu, wenn es keine Treffer in Freebase gibt.

Falls der Tag-Bezeichner nicht in Freebase und nicht in Wikipedia gefunden wird, werden anschließend WordNet und BeoLingus zur Bestimmung der Wortart abgefragt. Entsprechend der Annahme, die auch in [16] getroffen wird, werden Substantive dem Tag-Typ Thema und die anderen Wortarten keinem Tag-

¹⁰ http://www.webgenrewiki.org/index.php5/Genre_Classes_List, Zuletzt abgerufen am 28.02.2011

Typ zugeordnet, z.B. "schön" und "lustig" werden nicht als Substantive identifiziert und damit auch nicht als Thema. Bei Tags, die aus mehreren Wörtern bestehen, entscheidet die Anzahl der Vorkommen einer Wortart über den Tag-Typ, z.B. besteht "Internet Plattform" aus zwei Substantiven und wird damit als Thema mit dem Rang 102 klassifiziert.

Sortieren (8)

Der Sortieren-Schritt erhält eine Liste mit allen Tag-Typen der verschiedenen Klassifizierer, die zugeordnet werden konnten. Diese Liste wird anhand des Ranges sortiert, d.h. je niedriger der Rang, desto höher die Wahrscheinlichkeit für diesen Tag-Typ. Die Sortierung basiert wie oben beschrieben auf der Freebase-Reihenfolge oder den Heuristiken, denn für Tag-Typen, die nicht anhand von Freebase ermittelt wurden, werden die Tag-Typen anhand der Heuristik sortiert, die den Tag-Typ identifiziert hat. Beispielsweise werden dadurch Personen, die durch die Heuristik "[Anfangsbuchstabe + Name]" erkannt wurden, vorn in die Liste einsortiert, während Themen, die nur auf Grund der Wortart als Thema klassifiziert wurden, hinten in die Liste einsortiert werden. Der Tag-Typ an erster Stelle der sortierten Liste wird als Ergebnis ausgegeben. Konnte von den Klassifizierern kein Tag-Typ ermittelt werden, bleibt das Tag unkategorisiert.

5.3.2 Evaluation des eigenen Ansatzes

Das in dieser Arbeit entwickelte Verfahren wurde mit zwei verschiedenen Korpora evaluiert. Es wurde ein Korpus mit Delicious-Daten aufgebaut, der zum Vergleich mit dem Korpus dient, der aus den Daten der Evaluation des in dieser Arbeit implementierten Werkzeugs besteht.

Für die beiden Korpora wird deren Aufbau erläutert, anschließend werden die Ergebnisse der automatischen Erkennung der Tag-Typen beschrieben und diskutiert. Zum Abschluss wird ein Ausblick auf mögliche, zukünftige Verbesserungen des Verfahrens gegeben, die außerhalb dieser Arbeit liegen.

Erkennung von Tag-Typen in Delicious-Korpus

In diesem Abschnitt wird beschrieben, aus welchen Daten der Delicious-Korpus besteht, und wie gut der Ansatz diesen Korpus klassifiziert.

Aufbau des Korpus

Da ein entsprechender Vergleichskorpus nicht zugänglich ist, wurde ein eigener Korpus aufgebaut. Der Korpus ist zusammengestellt aus Bookmarks und Tags von Delicious. Bei der Zusammenstellung des Korpus war es wichtig, dass Bookmarks verschiedener Benutzer ausgewählt wurden, um möglichst viele verschiedene Tags im Korpus zu haben. Delicious listet kürzlich gespeicherte Bookmarks aller Nutzer auf einer speziellen Webseite auf¹¹, d.h. diese Liste enthält die Bookmarks verschiedener Nutzer.

Des Weiteren war es wichtig, dass der Korpus nicht nur Themen, sondern alle Tag-Typen enthält. Die Liste von Delicious kann anhand von Tags gefiltert werden. Empirisch wurde festgestellt, dass Ressourcen mit gewissen Tags gleichzeitig auch mit Tags bestimmter Tag-Typen getaggt sind. Die Tags "feiertag" und "urlaub" kommen beispielsweise häufig mit Orten, Zeitangaben und Ereignissen vor, die Tags "krieg",

¹¹ <http://www.delicious.com/recent/>, Zuletzt abgerufen am 28.11.2010

"spd" und "cdu" mit Orten und auch Personen bzw. Organisationen. Insgesamt wurden durch Filterung der Liste mit diesen Tags 225 Bookmarks mit 1272 Tags für den Korpus ausgewählt.

Die Wahrscheinlichkeit ist hoch, dass eine Ressource, die mit einem deutschsprachigen Tag getaggt wurde, mit einem weiteren deutschsprachigen Tag getaggt ist. Daher wurden deutschsprachige Tags zum Filtern der Liste in Delicious verwendet, um die Anzahl deutschsprachiger Delicious-Tags zu erhöhen, da die Benutzer des im Rahmen der Arbeit entwickelten Systems ELWMS.KOM in der Mehrheit ebenfalls deutschsprachig sind, und dieser Delicious-Korpus als Vergleichkorpus zum ELWMS.KOM-Korpus dienen soll.

Drei verschiedene Personen haben diese Tags von Hand mit den Tag-Typen ausgezeichnet. Das Fleiss' Kappa_{free} [114] beträgt 0,66 und entspricht laut Tabelle 5.1 einer beachtlichen Übereinstimmung.

Dieser Wert zeigt aber auch, dass die Zuordnung von Tags zu Tag-Typen nicht immer eindeutig möglich ist. Beispielsweise wurde das Tag "urlaub" manchmal als Thema und manchmal als Ereignis klassifiziert. Unstimmigkeiten gab es auch bei Tags, die im Grunde keine Themen sind, aber als Thema klassifiziert wurden, weil die Ressource dieses Tag beschreibt. Beispielsweise wurden Tags wie "Twitter" – eigentlich eine Organisation – als Thema klassifiziert, weil die getaggte Ressource über Twitter berichtet. Diesen Unterschied vollzieht das in dieser Arbeit entwickelte Verfahren nicht.

Ergebnisse

Die Ergebnisse eines automatischen Klassifizierers werden häufig in Konfusionsmatrizen dargestellt. Eine solche Matrix listet in der Diagonale die korrekt klassifizierten Objekte auf. Die anderen Werte geben die Anzahl an False Negatives (Zeile) und False Positives (Spalte) an. Je kleiner diese Werte sind, desto besser schneidet das Verfahren ab.

a	b	c	d	e	f	← klassifiziert als
91	2	0	0	6	7	a = Person
2	132	0	0	8	2	b = Ort
0	0	15	0	0	1	c = Ereignis
1	0	1	8	9	8	d = Typ
15	16	10	1	231	15	e = Thema
4	0	0	0	10	25	f = Unkategorisiert
0,81	0,88	0,58	0,89	0,88	0,43	Precision
0,86	0,92	0,94	0,30	0,80	0,64	Recall
0,83	0,90	0,71	0,44	0,84	0,52	F-Measure
81,0%						F-Measure Ø

Tabelle 5.5: Konfusionsmatrix für die Klassifikation der Tag-Typen im Delicious-Korpus

In Tabelle 5.5 ist die Konfusionsmatrix dargestellt für die Klassifikation der Tags, die übereinstimmend von allen drei Personen einem Tag-Typ zugeordnet wurden. Beispielsweise wurden 91 Tags korrekt als Person klassifiziert, zwei als Ort, sechs als Thema und sieben blieben unkategorisiert. Außerdem wurden z.B. zwei Orte, ein Typ, 15 Themen und vier ohne Tag-Typ falsch als Person klassifiziert, mehr zu den falsch klassifizierten Tags im Abschnitt 5.3.2.

Die Werte für Precision und Recall jeder Kategorie sind unterhalb der Matrix angegeben. Das F-Measure gewichtet über alle Kategorien beträgt 81%. 502 der 620 Tags (d.h. 81%) wurden korrekt klassifiziert.

a	b	c	d	e	f	← klassifiziert als
108	4	0	0	15	14	a = Person
3	146	0	0	8	2	b = Ort
0	0	20	0	3	2	c = Ereignis
2	0	2	19	26	9	d = Typ
44	25	14	11	387	38	e = Thema
22	3	5	1	33	52	f = Unkategorisiert
0,60	0,82	0,49	0,61	0,82	0,44	Precision
0,77	0,92	0,80	0,33	0,75	0,45	Recall
0,68	0,87	0,61	0,43	0,78	0,45	F-Measure
71,7%						F-Measure Ø

Tabelle 5.6: Konfusionsmatrix für die Klassifikation der Tag-Typen im Delicious-Korpus (2/3-Mehrheit)

In Tabelle 5.6 ist die Konfusionsmatrix dargestellt für Tags, die übereinstimmend von zwei der drei Personen einem Tag-Typ zugeordnet wurden. Insgesamt wurden 732 von 1018 Tags, d.h. 71,9%, korrekt klassifiziert.

Es wurde ebenso ermittelt, wie oft das Klassifikationsverfahren den Tag-Typ trifft, den die dritte Person angegeben hat, falls das Verfahren das Tag nicht dem Tag-Typ der Mehrheit zugeordnet hat, z.B. wenn zwei der drei Personen ein Tag als Typ klassifiziert haben, während die dritte Person dieses Tag als Thema klassifiziert hat. Der Tag-Typ der Mehrheit und der dritten Person werden beide als korrekt gewertet. Die Ergebnisse sind in Tabelle A.5 im Anhang dargestellt, das gewichtete F-Measure beträgt 78,2%.

Diskussion

Die meisten Fehler geschehen bei der Entscheidung zwischen dem Tag-Typ Thema und den anderen Tag-Typen. Manche dieser Fehlentscheidungen resultieren aus der durch die Ersteller des Korpus erfolgten Zuordnung zu Themen, wenn die Ressource von dem Tag handelt (z.B. Twitter als Thema der Ressource statt Person). Der Algorithmus kann solche Tags nicht automatisch als Thema klassifizieren.

Des Weiteren gibt es Tags, die falsch geschrieben wurden, und daher nicht in den externen Quellen gefunden werden. Bei der menschlichen Einordnung wurden falsch geschriebene Tags dennoch zu den Tag-Typen zugeordnet. Diese menschliche Leistung kann der automatische Ansatz so nicht leisten, mehr dazu im Abschnitt 5.3.2. Zusätzlich enthält der Korpus Fehler. Beispielsweise wurden Tags nicht als Orte klassifiziert, weil sie den menschlichen Klassifizierern unbekannt sind.

Neben diesen Fehlklassifikationen erhöht sich die Fehlerwahrscheinlichkeit durch die Übersetzungen der Tags. Dieser Schritt entfällt in den verwandten Ansätzen, da diese nicht-englischsprachige Tags nicht klassifizieren. Teilweise entsteht auch durch die Abfrage mit deutschen Termen eine Mehrdeutigkeit. Verschiedene deutsche, allgemeine Tags sind beispielsweise auch Ortsbezeichner und werden beispielsweise aufgrund ihres hohen Rangs in Freebase als Orte klassifiziert z.B. wetter, kinder, heiden, tannenbaum, trojaner.

Daher wurden aus diesem Korpus manuell die nicht-englischsprachigen Tags aussortiert, um das Verfahren nur auf Eigennamen und englisch-sprachige Tags anzuwenden. Die Ergebnisse der Klassifikation sind in Tabelle A.1 und A.2 im Anhang aufgelistet, das F-Measure erhöht sich dabei auf bis zu 86,6%.

Im Folgenden sollen konkrete Beispiele die Fehlklassifikationen verdeutlichen.

- taj mahal: Ort, der falsch als Person klassifiziert wurde, da der Eintrag in Freebase zur Person höher gerankt ist als der Ort.
- dossier: Typ, der falsch als Person (Farewell Dossier) klassifiziert wurde.
- datteln: Ort, der beim Aufbau des Korpus als Thema klassifiziert wurde, und deshalb als Fehlklassifikation zählt.

Evaluation mit typisierten Tags aus ELWMS.KOM

In diesem Abschnitt wird beschrieben, wie gut der automatische Ansatz für Tags in ELWMS.KOM funktioniert.

Aufbau des Korpus

Zum Zeitpunkt des Korpus-Aufbaus gab es in ELWMS.KOM 422 Ressourcen und 1161 Tags. Diese wurden durch die Nutzer von ELWMS.KOM bereits typisiert. Da bekannt ist, dass die Typisierung teilweise subjektiv ist, wurden die Tags von zwei weiteren Personen klassifiziert, so dass es insgesamt für jedes Tag drei Typen gibt. Den Personen wurde bei der Klassifizierung der vom Nutzer von ELWMS.KOM bereits definierte Typ nicht gezeigt. Das Fleiss' Kappa_{free} [114] beträgt 0,57 für alle Tags und entspricht laut Tabelle 5.1 nur einer mittelmäßigen Übereinstimmung. Sehr häufig klassifizieren die beiden Personen von den Nutzern von ELWMS.KOM erstellte Ziele übereinstimmend als Themen. Es gibt insgesamt 366 Ziel-Tags, die 174 mal von beiden Personen als Thema klassifiziert wurden. Nur 65 Ziel-Tags wurden übereinstimmend als Ziel klassifiziert. Werden Ziele außer Acht gelassen, beträgt das Fleiss' Kappa_{free} 0,67 und entspricht damit einer beachtlichen Übereinstimmung.

Ergebnisse

Die Ergebnisse des Verfahrens auf dem Korpus von ELWMS.KOM werden anhand der Konfusionsmatrizen dargestellt.

Tabelle 5.7 zeigt die Ergebnisse, die erreicht wurden, für die Klassifikation der Tags, die übereinstimmend von allen drei Personen einem Tag-Typ zugeordnet wurden. Tags vom Tag-Typ Ziel wurden hierbei nicht klassifiziert. Das F-Measure gewichtet über alle Kategorien beträgt 80,9%. 371 der 493 Tags (d.h. 75,3%) wurden korrekt klassifiziert.

Tabelle 5.8 zeigt die Ergebnisse für die Klassifikation der Tags, die übereinstimmend von zwei der drei Personen einem Tag-Typ zugeordnet wurden. Das F-Measure gewichtet über alle Kategorien beträgt 71,5%. 639 der 948 Tags (d.h. 67,4%) wurden korrekt klassifiziert.

Es wurde ebenso ermittelt, wie oft das Klassifikationsverfahren den Tag-Typ trifft, den die dritte Person angegeben hat, falls das Verfahren das Tag nicht dem Tag-Typ der Mehrheit zugeordnet hat. Der Tag-Typ der Mehrheit und der dritten Person werden beide als korrekt gewertet. Die Ergebnisse sind in Tabelle A.6 im Anhang dargestellt, das gewichtete F-Measure beträgt 74,5%.

a	b	c	d	e	f	← klassifiziert als
140	1	0	0	2	25	a = Person
0	14	0	0	2	0	b = Ort
0	0	53	0	3	0	c = Ereignis
0	5	0	4	10	0	d = Typ
18	2	1	0	160	53	e = Thema
0	0	0	0	0	0	f = Unkategorisiert
0,89	0,63	0,98	1,00	0,90	0,00	Precision
0,83	0,88	0,95	0,21	0,68	0,00	Recall
0,86	0,74	0,96	0,35	0,78	0,00	F-Measure
80,9%						F-Measure Ø

Tabelle 5.7: Konfusionsmatrix für die Klassifikation der Tag-Typen in ELWMS.KOM

a	b	c	d	e	f	← klassifiziert als
170	3	1	0	12	31	a = Person
1	15	0	0	3	0	b = Ort
6	0	62	0	4	1	c = Ereignis
0	5	0	6	53	0	d = Typ
36	7	11	4	380	129	e = Thema
0	0	0	0	2	6	f = Unkategorisiert
0,80	0,50	0,84	0,60	0,84	0,04	Precision
0,78	0,79	0,85	0,09	0,67	0,75	Recall
0,79	0,61	0,84	0,16	0,74	0,07	F-Measure
71,5%						F-Measure Ø

Tabelle 5.8: Konfusionsmatrix für die Klassifikation der Tag-Typen in ELWMS.KOM (2/3-Mehrheit)

Aus diesem Korpus wurden auch manuell die nicht-englischsprachigen Tags aussortiert, um das Verfahren nur auf Eigennamen und englischsprachige Tags anzuwenden. Die Ergebnisse der Klassifikation sind in Tabelle A.3 und A.4 im Anhang aufgelistet, das F-Measure erhöht sich dabei auf bis zu 82,5%.

Diskussion

Die meisten Fehler geschehen auch hier bei der Entscheidung zwischen dem Tag-Typ Thema und den anderen Tag-Typen. Die in Abschnitt 5.3.2 genannten Ursachen für Fehlentscheidungen (z.B. deutschsprachige Tags, Thema der Ressource etc.) treffen auch für diesen Korpus zu.

Zusätzlich gibt es in den Korpusdaten viele Personen, für die es keine Einträge in den externen Quellen gibt. Die Personendatenbank enthält Vor- und Familiennamen aus Deutschland und den USA. Aufgrund der Namen der zu klassifizierenden Personen, die weder typisch deutsch noch amerikanisch sind, können diese Personen nicht klassifiziert werden; das betrifft beispielsweise Namen aus dem asiatischen Raum.

Die Liste der Webgenres enthält viele der Typen nicht, weshalb der Recall für Typen sehr gering ist. Bei einem Einsatz des Verfahrens muss die Liste der Webgenres noch erweitert werden.

Des Weiteren gibt es viele Themen, die aus verschiedenen Fachtermini zusammengesetzt sind. Diese können in der zusammengesetzten Form weder in Wikipedia gefunden werden noch kann die Wortart bestimmt werden.

Im Folgenden sollen konkrete Beispiele die Fehlklassifikationen verdeutlichen.

- Interactive graphics for Data Analysis: Aus verschiedenen Fachwörtern zusammengesetztes Thema, das nicht in einer externen Quelle gefunden werden kann.
- "28" und "September": Sind zwei einzelne Tags, die die menschlichen Klassifizierer allerdings zu einem Tag zusammengefasst und dann die Zusammensetzung als Ereignis klassifiziert haben. Das automatische Verfahren fasst keine Tags zusammen.
- MP2P: Abkürzung eines Fachterminus, der nicht in den externen Quellen gefunden wurde.

Weitere Aspekte des Verfahrens

Neben dem gezeigten Basisverfahren wurden verschiedene weitere Aspekte im Rahmen dieser Arbeit untersucht, die im Folgenden beschrieben werden.

Laufzeitverhalten

Die Dauer der Klassifikation eines Tags liegt im Durchschnitt unter 1 Sekunde je Tag (für Informationen zur Messung der Klassifikationszeit siehe Anhang A.7). Dies ist ausreichend schnell, um dem Benutzer bereits während der Eingabe des Tags einen Tag-Typ vorzuschlagen.

Alternativ können die Tags auch in Zeiten der Inaktivität, z.B. über Nacht, automatisch klassifiziert werden und dem Benutzer bei seinem nächsten Login oder während einer speziellen Bearbeitungsfunktion der Tagging-Struktur präsentiert werden, in der der Benutzer die Fehler des Verfahrens korrigieren kann.

Annahme zur URL der Ressource

Im Ansatz von Bischoff et al. [16] wurde die Annahme getroffen, dass das Enthaltensein des Tag-Bezeichners in der URL der getaggten Ressource auf den Tag-Typ Person/Organisation hinweist. Daher

wurde das Verfahren um diese Annahme erweitert. In Abbildung A.1 ist das Vorgehen zur Identifikation des Tag-Typs schematisch dargestellt, bei dem die URL der Ressource mit einbezogen wird. Eingabeparameter sind neben dem vom Nutzer eingegebenen Tag-Bezeichner somit auch die URL der getaggten Ressource. Ist der Tag-Bezeichner in der Domain der URL enthalten, wird dieses Tag als Person/Organisation klassifiziert, z.B. die Ressource "http://www.kom.tu-darmstadt.de" getaggt mit "kom" würde dazu führen, dass das Tag "kom" als Person klassifiziert wird. In Tabellen A.9 und A.10 im Anhang sind die Ergebnisse des Verfahrens unter Berücksichtigung der URL für den Delicious-Korpus dargestellt. Zwar erhöht sich bei diesem Korpus das F-Measure von 81% auf 81,5% und von 71,7% leicht auf 72,1% für die 2/3-Mehrheit, da durch diese Annahme mehr Tags korrekt als Personen klassifiziert werden können. Allerdings werden dadurch auch einige der anderen Tag-Typen falsch als Person klassifiziert, z.B. erhöht sich die Anzahl der falsch als Person klassifizierten Themen von 44 auf 53.

Im ELWMS-Korpus erhöht sich das F-Measure von 80,9% auf 81,3% und sinkt von 71,5% leicht auf 71,3% für die 2/3-Mehrheit, da auch in diesem Korpus zwar mehr Personen korrekt klassifiziert werden können, aber es auch mehr Falschklassifikationen gibt (siehe Tabelle A.11 und A.12 im Anhang).

DBpedia als zusätzliche Wissensquelle

Neben Freebase sind weitere Wissensquellen wie DBpedia denkbar, um das Verfahren zu optimieren und weitere bisher nicht kategorisierte Tags klassifizieren zu können. Schematisch ist der Ablauf des Verfahrens unter Hinzunahme von DBpedia in Abbildung A.2 im Anhang dargestellt. Wird der Tag-Bezeichner nicht in Freebase gefunden und trifft keine der Heuristiken für das Tag zu, wird anschließend noch DBpedia abgefragt (Schritt 9). Die DBpedia-Kategorien lassen sich auf die Tag-Typen Person, Ereignis und Ort abbilden. Falls DBpedia einen Treffer zum Suchwort zurückliefert, wird dessen Kategorien ermittelt. Kann der Treffer keinem der drei Tag-Typen zugeordnet werden, wird er dem Tag-Typ Thema zugeordnet, denn durch einen existierenden Eintrag in DBpedia wird angenommen, dass der Tag-Bezeichner eine nutzerübergreifende Bedeutung hat. Durch die zusätzliche Wissensquelle können mit diesem Schritt weitere bisher nicht-kategorisierte Tags identifiziert werden. Beispielsweise wird im Orts-Klassifizierer der Ort "Gardasee" nicht durch GeoNames und Freebase als Ort identifiziert, weil GeoNames in diesem Fall keinen Treffer zurückliefert. Durch DBpedia kann der Gardasee aber als Ort klassifiziert werden. Die Ergebnisse sind in den Tabellen A.13 und A.14 für Delicious und A.15 und A.16 für ELWMS.KOM aufgelistet. Das F-Measure beträgt bei der 2/3-Mehrheit bis zu 74% und bei der übereinstimmenden Mehrheit bis zu 83,7%.

Die Ergebnisse für das Verfahren unter Berücksichtigung der URL und gleichzeitiger Hinzunahme von DBpedia sind im Anhang im Abschnitt A.3 aufgelistet. Das F-Measure erhöht sich auf bis zu 84,2%. Es gibt allerdings auch eine erhöhte Anzahl an Falsch-Klassifikationen.

Das Hinzunehmen einer weiteren Datenquelle hat verschiedene Nachteile: Zum einen wird das Verfahren von einer weiteren Datenquelle abhängig und somit erhöht sich der Pflegeaufwand für das Verfahren. Zum anderen erhöht sich dadurch die Laufzeit des Verfahrens, also die Zeit, die für die Klassifikation eines Tags benötigt wird. Des Weiteren erhöht sich die Wahrscheinlichkeit für Falschklassifikationen.

Erkennung der thematischen Ziele

Ziel des entwickelten Verfahrens ist die Erkennung von Themen, Orten, Personen, Ereignissen und Typen. Der Tag-Typ Ziel blieb bisher unberücksichtigt, da es in ELWMS.KOM zur Erstellung von Zielen

einen speziellen Mechanismus gibt (siehe Kapitel 8.2.1) und Ziele daher nicht automatisch klassifiziert werden müssen. Bei der Bestimmung des Tag-Typs beim Aufbau des ELWMS-Korpus wurden allerdings viele der ursprünglichen Ziele von den beiden Personen übereinstimmend als Thema klassifiziert. Solche thematischen Ziele sind, wie in Kapitel 4.4 erwähnt, für das Empfehlen von Tags und Ressourcen interessant. Daher wurde ermittelt, wie gut das Verfahren thematische Ziele erkennen kann. Dazu wurde der Korpus auf die Tags reduziert, die vom Nutzer von ELWMS.KOM ursprünglich als Ziel erstellt wurden und von den beiden anderen Personen einem der anderen Tag-Typen zugeordnet wurden. In Tabelle A.7 im Anhang ist die Konfusionsmatrix dargestellt, es werden beispielsweise 137 der thematischen Ziele vom Verfahren dem Tag-Typ Thema zugeordnet, 28 bleiben unkategorisiert. Insgesamt kann auf diesem Teilkorpus ein F-Measure von 70,8% erreicht werden.

Werden nicht nur die thematischen Ziele, sondern alle Ziele des Nutzers vom Verfahren klassifiziert, dann reduziert sich das F-Measure auf 55,4% (siehe Tabelle A.8 im Anhang). Der Grund für das eher schlechte Ergebnis liegt u.a. darin, dass es keinen speziellen Klassifizierer für Ziele gibt. Ziele enthalten häufig Substantive, so dass das Verfahren diese dem Tag-Typ Thema zuordnet und zwischen Ziel und Thema somit nicht gut unterscheiden kann. Das Verfahren kann daher den Nutzer von ELWMS.KOM zwar beim Finden von thematischen Zielen semi-automatisch unterstützen; für eine rein automatische Bestimmung von thematischen Zielen muss aber zukünftig ein spezieller Ziel-Klassifizierer ergänzt werden, beispielsweise weist das Vorhandensein von Verben im Tag-Bezeichner auf den Tag-Typ Ziel hin.

5.4 Zusammenfassung und zukünftige Arbeiten

In diesem Kapitel wurden die Grundlagen für ein Verfahren zur automatischen Erkennung von Tag-Typen erläutert und es wurden verwandte Ansätze beschrieben. Des Weiteren wurden das in dieser Arbeit entwickelte Verfahren und dessen Evaluation beschrieben. Das Verfahren wurde mit zwei verschiedenen Korpora evaluiert. Auf diesen Korpora erreicht das in dieser Arbeit entwickelte Verfahren mit einer Genauigkeit von bis zu ca. 84% vergleichbar gute Ergebnisse. Auf Grund des wissensbasierten Ansatzes statt eines korpusbasierten Verfahrens wird nur eine geringe Dauer für die Identifikation des Tag-Typs benötigt, so dass sich der Tag-Typ bereits bei Eingabe des Tags bestimmen lässt. Fälle, in denen das Verfahren nicht korrekt arbeitet, wurden im entsprechenden Abschnitt der beiden Korpora erläutert. Extrahierte Listen wurden für das Verfahren nicht manuell angepasst und ergänzt, damit diese Listen nicht auf den jeweiligen Korpus abgestimmt sind, auch wenn die Ergebnisse durch Ergänzung beispielsweise von weiteren Webgenres verbessert werden könnten.

Das Verfahren hat noch Potential für Verbesserungen. Beispielsweise kann versucht werden, Tippfehler automatisch zu erkennen, dabei ist zu beachten, dass bei Annahme eines Tippfehlers aus dem Wort "Reise" schnell "Preise" werden kann. Auch ist zu prüfen, inwieweit Tippfehler bei Ortsnamen und Personen erkannt werden können.

Für die Unterscheidung der Bedeutung von Tags gibt es Arbeiten im Bereich der Tag Disambiguation, die auf Basis von häufig gleichzeitig vorkommenden Tags (Co-Occurrences) oder dem Inhalt der Ressource die Bedeutung unterscheiden. Problematisch hierbei ist, dass damit nur Tags erkannt werden können, deren Bedeutung Inhalt bzw. Kontext der Ressource ist, d.h. es wird eine bestimmte Menge von Co-Occurrences bzw. eine bestimmte Menge an Ressourcen-Inhalt benötigt. In [44] wird beispielsweise ein Verfahren beschrieben, das die Bedeutung von Tags mittels DBpedia unterscheiden will.

Wie erläutert, können die Tag-Typen Person bzw. Organisation, Ereignis, Typ und Ort abhängig von der Ressource manchmal als Thema der Ressource klassifiziert werden. Denn eine Ressource kann beispielsweise über eine Person berichten. Diese Unterscheidung wird in der Basismenge der Tag-Typen nicht vollzogen, könnte aber bei entsprechender Erweiterung des automatischen Verfahrens durch Facetten realisiert werden. Es könnte dann z.B. die Facette "Person" – "als Thema der Ressource" geben. Dazu muss der Kontext des Tags, also der Inhalt der Ressource, einbezogen werden.



6 Bestehende Ansätze und Werkzeuge zur Unterstützung des Ressourcen-basierten Lernens

Es existieren verschiedene Ansätze und Werkzeuge, die teilweise die Anforderungen des Ressourcenmanagements (siehe Kapitel 2.3) erfüllen. Dazu zählen Conceptual Diagrams, Semantic Desktops, Persönliche Lernumgebungen, Social Bookmarking-Dienste und Werkzeuge zum Zielmanagement. Im Folgenden sollen diese verwandten Ansätze erläutert und anhand von ausgewählten Beispielen veranschaulicht werden. Am Ende dieses Kapitels wird zusammenfassend diskutiert, inwieweit jeder dieser Ansätze die Prozessschritte des Ressourcenmanagements (siehe Modell in Kapitel 2.3.3) unterstützt.

6.1 Conceptual Diagram

Conceptual Diagrams [164] sind graphische Beschreibungen von Konzepten und deren Beziehungen. Die graphische Darstellung kann mittels Kreisen, Kästchen, Pfeilen, Pyramiden, usw. erfolgen. Software-basierte Conceptual Diagrams ermöglichen nicht nur die Visualisierung der Struktur der Konzepte, sondern sie erlauben auch die Einbindung von Ressourcen z.B. als Links zum Web und die Erstellung von Annotationen innerhalb des Conceptual Diagrams. Weiterer Vorteil der Software-basierten Conceptual Diagrams ist die Suche in den Ressourcen. Beispiele für Conceptual Diagrams sind Mind-Maps und Concept-Maps.

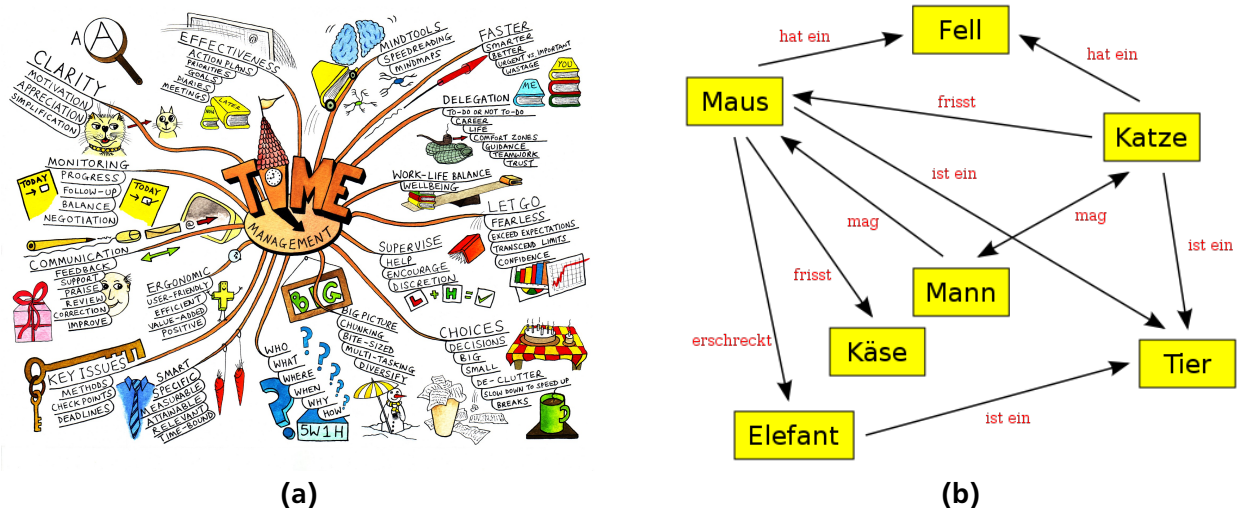


Abbildung 6.1: a) Mind-Map¹, b) Concept-Map²

Mind-Maps

Mind-Maps [24] werden vor allem zum Festhalten erster Gedanken und Ideen eingesetzt. Sie bestehen aus einem zentralen Hauptgedanken in der Mitte, zu dem nach und nach radial weitere Ideen hinzuge-

¹ <http://www.mindtools.com/media/Diagrams/mindmap.jpg>, Zuletzt abgerufen am 09.01.2011

² http://de.wikipedia.org/w/index.php?title=Datei:Concept-Map_de.svg&filetimestamp=20081223211452, Zuletzt abgerufen am 09.01.2011

fügt werden, die astförmig von der Mitte zum Rand verlaufen. Sie können (sehr) farbig gestaltet sein und Bilder enthalten. In Abbildung 6.1a wird eine bilderreiche Mind-Map gezeigt.

Concept-Maps

Concept-Maps [25] werden eingesetzt, um Wissen in Form von Konzepten und deren Beziehungen untereinander darzustellen. Im Allgemeinen werden die Konzepte in Kreisen oder Kästchen und die Beziehungen durch Linien oder gerichtete Pfeile visualisiert. Concept-Maps können hierarchisch aufgebaut sein, dabei werden die zentraleren Konzepte weiter oben angeordnet. Die Hierarchie ist allerdings nicht zwingend. Abbildung 6.1b zeigt ein Beispiel für eine Concept-Map.

Diskussion

Die Effekte des Einsatzes von Mind-Maps oder Concept-Maps auf den Lernprozess und Lernerfolg wurden in zahlreichen Artikeln diskutiert, u.a. [35, 157, 41, 164]. In [97] werden mehr als 500 Arbeiten identifiziert, die den Einsatz in Lernszenarien untersuchen. Neben dem Einsatz beim Lernen werden die Maps auch zur Navigation auf Webseiten eingesetzt [122]. In [13] wird vorgeschlagen, bereits vom Nutzer erstellte Mind-Maps zur Verbesserung von Suchergebnissen zu nutzen, indem die Beziehungen zwischen benachbarten Konzepten beim Ranking oder Filtern genutzt werden.

Mind-Mapping Techniken sind leicht zu erlernen, da die radiale Struktur fest vorgegeben ist. Das Erstellen von Concept-Maps benötigt eine gewisse Einarbeitung, da sie sehr komplex werden können. In [40] werden drei Einsatzbereiche genannt: Gliederung von Lerninhalten zur Vorbereitung, Wissenstransfer bei gemeinsamer Arbeit und Erarbeitung von Bedeutungszusammenhängen. Wobei Mind-Maps häufiger eingesetzt werden, um kreative Ideen in Brainstormings festzuhalten und Concept-Maps zur Aufbereitung von Lerninhalten.

Die Ergebnisse der Untersuchungen in [98] zeigen, dass Mind-Mapping vor allem die Prozesse der Wissensidentifikation und Wissensbewertung unterstützt. Die Lernenden können ihr Vorwissen aktivieren und dabei Wissenslücken identifizieren. Im Laufe des weiteren Lernprozesses kann die Mind-Map zur Kontrolle der Lernfortschritte dienen. Conceptual Diagrams können die Generierung, Strukturierung, Organisation und Repräsentation von Wissen und Informationen erleichtern [71]. Neues Wissen entsteht vor allem dort, wo neue Konzepte hinzugefügt und verbunden werden. In [25] wird gezeigt, dass Concept-Maps eine große Bedeutung bei der Visualisierung und Kommunikation von Wissen haben.

Conceptual Diagrams können nur eine geringe Menge an Konzepten übersichtlich darstellen. Verbesserungen werden daher entwickelt, die beispielsweise die Gruppierung von Konzepten ermöglichen und Verbindungslinien bei Bedarf ein- bzw. ausblenden [55].

6.2 Semantic Desktop

Ein Semantic Desktop wird in [124] definiert als *"a device in which an individual stores all her digital information such as documents, multimedia and messages. [...] Semantic Web protocols are used for inter-application communication. The use of Semantic Web standards allows existing web resources to be in-*

corporated into the personal knowledge space, and does also facilitate the sharing of knowledge with others [...]".

Ziel des Semantic Desktop ist laut dieser Definition die Nutzung von Technologien des Semantic Web auf dem Desktop von Endanwendern zum Zweck des persönlichen Informationsmanagements. Die Organisation der Informationen (Dokumente, Multimedia und Nachrichten) soll dabei über Applikationsgrenzen hinweg ermöglicht werden. Die Verwendung von Semantic Web-Standards ermöglicht den Austausch der Informationen mit anderen Semantic Desktops.

Es gibt verschiedene Ansätze, die bestimmte Aspekte des Semantic Desktops fokussieren, beispielsweise Gnowsis (Integration in existierende Desktop-Anwendungen) [124], Haystack (integrierte Umgebung ohne Applikationsgrenzen) [109], D-BIN (Semantic Desktop auf Basis von Peer-to-Peer-Netzen) [150], seMouse (semantische Annotation in Dokumenten) [70] und Chandler³ (Austausch von Informationen). Ein Überblick über diese und weitere Ansätze findet sich in [124]. Einige der existierenden Ansätze wurden im EU-Projekt Nepomuk⁴ integriert und weiterentwickelt. Ziel dieses Projektes war, die Vernetzung der Semantic Desktops zu einem *Social Semantic Desktop* weiter zu entwickeln. Die Vision des Social Semantic Desktop wurde bereits in [37] beschrieben und wird manchmal auch als *Networked Semantic Desktop* bezeichnet.

Am Beispiel von Gnowsis wird im Folgenden der Social Semantic Desktop näher beschrieben. Dieser Ansatz wurde zur Beschreibung ausgewählt, weil er den Anforderungen des Anwendungsszenarios in dieser Arbeit am ehesten gerecht wird. Der Austausch der Informationen findet in RDF⁵ statt. RDF dient zur formalen Beschreibung von Informationen über Objekte, in RDF als Ressourcen bezeichnet. Ressourcen sind zum einen Webseiten, Dokumente oder Dateien und zum anderen Dinge der realen Welt. Diese Ressourcen werden durch eindeutige Bezeichner (Unified Resource Identifiers) identifiziert. Mit RDF lassen sich Sachverhalte (sogenannte Statements) wie Beziehungen zwischen Ressourcen und Eigenschaften von Ressourcen formal beschreiben. RDF Statements können in XML ausgedrückt werden, womit sich die Informationen zwischen Systemen austauschen lassen.

Die Basis für das persönliche Informationsmanagement in Gnowsis bildet das *persönliche Informationsmodell* (PIMO), das einer persönlichen Ontologie entspricht. Es ist die Formalisierung der subjektiven Sicht eines Benutzers, die dieser mit seinen Applikationen teilt. Die Informationen in den verschiedenen Applikationen werden in diesem Informationsmodell kumuliert abgebildet. Ressourcen können kategorisiert und mit Konzepten getaggt werden. Domänenunabhängige Klassen der Ontologie sind "Person", "Thema", "Dokument", "Ort" und "Ereignis". Das Informationsmodell kann durch den Benutzer oder domänenspezifische Ontologien erweitert werden und soll dem Benutzer die Externalisierung seines Wissens ermöglichen. Kapitel 4.3.1 geht näher auf die Basisklassen des Informationsmodells ein.

In Gnowsis wurde ein Service-orientierter Ansatz gewählt, so dass sich neue Applikationen (leicht) integrieren lassen, die das persönliche Informationsmodell und die Technologien des Semantic Desktop nutzen können. Zum Aufbau des persönlichen Informationsmodells sind in Gnowsis verschiedene Applikationen integriert, mit denen die Ressourcen semantisch angereichert werden können. Die "Miniquire"-Sidebar (Abbildung 6.2 links) gibt einen Überblick über das PIMO. Es kann durchsucht, gefiltert und editiert werden. Der "ThingEditor"-Browser (Abbildung 6.2 rechts) öffnet sich beim Editieren. Der Editor

³ <http://chandler.osafoundation.org/>, Zuletzt abgerufen am 19.01.2011

⁴ NEPOMUK - The Social Semantic Desktop - FP6-027705, <http://nepomuk.semanticdesktop.org>, Zuletzt abgerufen am 19.01.2011

⁵ Resource Description Framework (RDF) [115]

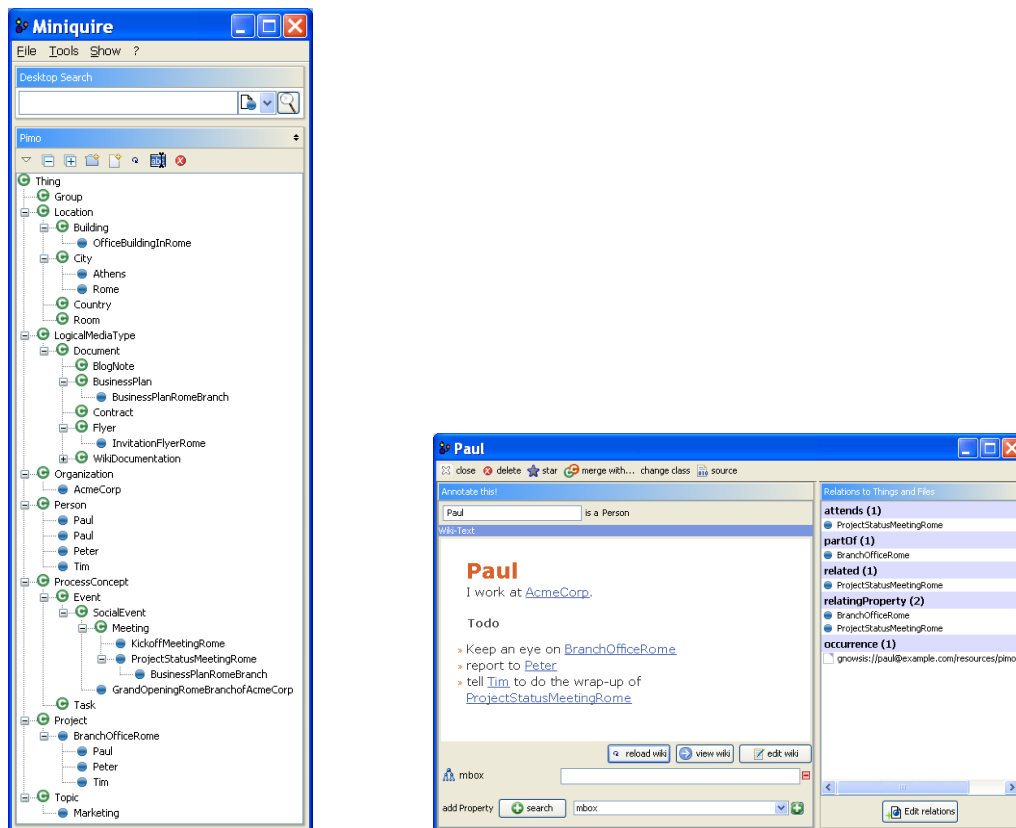


Abbildung 6.2: "Miniquire"-Sidebar (links) und "ThingEditor"-Browser (rechts) [124]

zeigt alle Relationen und Metadaten des geöffneten Objektes an. Andere Objekte können neu erstellt werden oder per Drag&Drop aus der Sidebar dem Objekt hinzugefügt werden. Ein freies Textfeld zeigt den Text aus dem Semantic Wiki an.

Als Semantic Wiki ist Kaukolu [74] integriert. Für jede Instanz im Informationsmodell kann eine Wiki-Seite erstellt werden. Bei der Bearbeitung des Wiki-Textes wird der Benutzer durch eine Auto-Vervollständigung für die PIMO-Instanzen unterstützt. Mit einer speziellen Semantic Wiki-Syntax können neue Konzepte und Relationen zum Modell hinzugefügt werden. Beispielsweise wird "Rom" durch den Ausdruck "[Rom] ist eine [Stadt]" als Instanz der Klasse "Stadt" erzeugt. Jede neue Wiki-Seite wird automatisch zum Tag in Gnowsis.

Das Tagging von Webseiten in Gnowsis wird durch ein Add-on im Webbrowser realisiert, die Funktionalität des Tagging-Dialogs ist ähnlich zu Delicious (siehe Abschnitt 6.4). Beim Tagging-Vorgang wird das zu taggende Dokument analysiert und auf Basis von Textähnlichkeiten werden Tags vorgeschlagen. Des Weiteren werden die Binärdaten des Dokuments in RDF gespeichert, so dass das Dokument später in Volltextsuchen zur Verfügung steht.

Es sind in Gnowsis weitere Anwendungen wie eine Timeline-Ansicht, eine Tag Cloud, eine Drop-Box etc. integriert.

Ziel des Semantic Desktop ist die Applikations-übergreifende Unterstützung der Organisation, des Wiederfindens, der Generierung und Kommunikation von Informationen und Wissen, das in Ressourcen, wie E-Mails, Kalendereinträgen, Dateien, Adressbucheinträgen etc. auf dem Desktop und im Web enthalten ist. Dabei werden Ressourcen mit semantischen Metadaten angereichert.

Semantic Desktops sind nicht für Lernszenarien konzipiert. Sie sind äußerst komplex und der Benutzer erhält wenig Unterstützung bei Aufbau und Pflege des persönlichen Informationsmodells. Durch Tagging von Webseiten wird das Modell zwar erweitert, aber es wird bei Suchen nach neuen Webseiten nicht genutzt, um im Web Orientierung zu bieten. Orientierungshilfen gibt es nur für bereits gespeicherte Dokumente.

Mit dem Social Semantic Desktop wurde eine technologische Basis geschaffen für eine Kollaboration und einen Austausch von Informationen. Bisher gibt es aber nur komplizierte Werkzeuge für Experten im Bereich des Ontologie-Mappings, um Ontologien miteinander auszutauschen. Mit diesen Werkzeugen ist der angestrebte Austausch von persönlichen Informationsmodellen für Laien nicht durchführbar.

6.3 Personal Learning Environment

Personal Learning Environments (PLEs) fokussieren auf die selbstgesteuert Lernenden. Bei einer PLE handelt es sich zunächst nicht um eine Software-Applikation, sondern um ein Konzept zur Nutzung von Technologien zum Lernen [7]. Es basiert auf einem Service-orientierten Ansatz zur Integration verschiedener Dienste. Basis der Integration bilden offene Standards und leichtgewichtige Schnittstellen zur Koordination zwischen Diensten und Nutzern der PLE.

Lernmanagementsysteme wurden mit dem Fokus auf einen Einsatz in Bildungsinstitutionen oder Organisationen entwickelt und beinhalten deshalb Funktionen zur Verwaltung von Kursen, Inhalten und Mitgliedern. Lernmanagementsysteme optimieren damit formales Lernen, während PLEs eher informelles Lernen, das beispielsweise nicht durch Studienordnungen eingeschränkt wird, unterstützen [132]. PLEs sind persönliche Lernumgebungen und benötigen diese Verwaltungsfunktionen daher nicht, denn die PLEs sind Eigentum der Lernenden. *Persönlich* bedeutet in diesem Zusammenhang aber auch personalisiert, d.h. die Lernenden können die Umgebung an ihre individuellen Wünsche und Bedürfnisse anpassen. In Lernmanagementsystemen ist die Anpassbarkeit durch die Lernenden stark eingeschränkt. Die Inhalte in PLEs können vor dem Zugriff von außen geschützt werden. In Lernmanagementsystemen haben Lehrende und Administratoren zumeist (uneingeschränkten) Zugriff auf die Inhalte und Daten. Der Lernende wird durch PLEs nicht an eine Institution gebunden, wenn es sich um eine Zusammenstellung freier Dienste handelt. Das technologische Konzept beschreibt Funktionen, die eine PLE haben sollte: Zusammenarbeit, Kommunikation und Vernetzung von Lernenden. Häufig werden Social Software-Anwendungen wie Wikis und Weblogs in die PLE eingebunden. Mit diesen Werkzeugen können Lernende nicht nur Inhalte konsumieren, sondern auch selbst erstellen und veröffentlichen.

Abbildung 6.3 beschreibt ein konzeptuelles Modell einer PLE [161]. Es werden verschiedene Dienste dargestellt, die mit der PLE kommunizieren. Beispielsweise können Inhalte von Weblogs über Schnittstel-

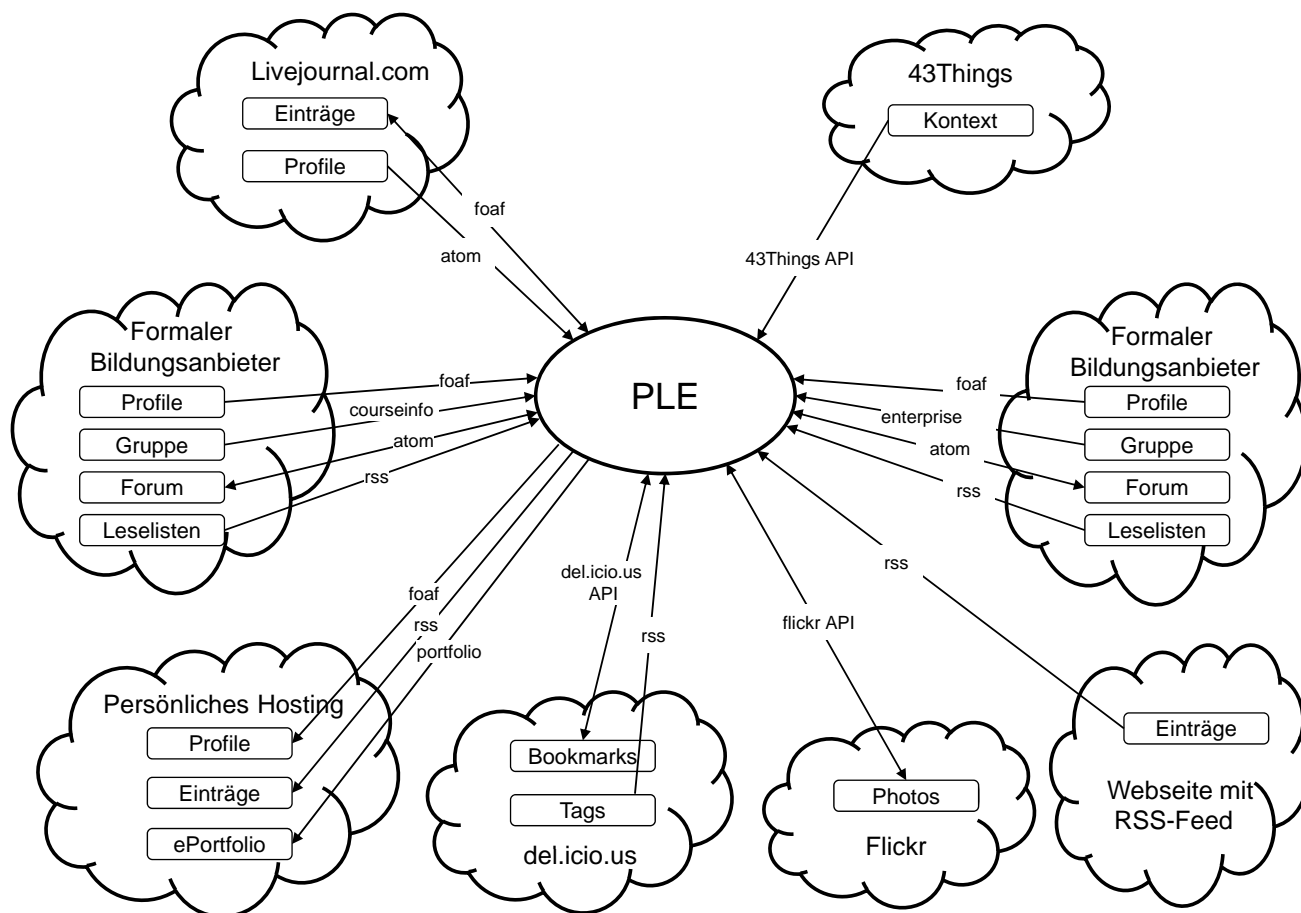


Abbildung 6.3: Konzeptuelles Modell einer Personal Learning Environment (vgl. [161])

len wie RSS⁶ eingebunden werden. Auch die Inhalte eines Lernmanagementsystems können integriert werden, wenn sie vom Lernmanagementsystem entsprechend bereitgestellt werden.

Realisierungen von PLEs sind beispielsweise Elgg⁷ und Flock⁸. Elgg ist eine Applikation mit dem Fokus auf der Erstellung eines sozialen Netzwerks und Lerngemeinschaften. Es enthält u.a. Werkzeuge zur Erstellung von Podcasts, Weblogs und Webseiten. Flock ist ein Browser, der auf die Integration verschiedener Arten von Social Software spezialisiert ist. Auch I-Google⁹ lässt sich als PLE einrichten. Dazu können Nutzer verschiedene Dienste wie Wörterbücher, Weblogs etc. auf einem persönlichen Portal zusammenstellen. Weitere PLEs werden u.a. in [151] beschrieben.

Diskussion

Die Werkzeuge in einer PLE lassen sich als Dienste entsprechend der Bedürfnisse des Lernenden zusammenstellen. Die PLE fungiert als eine Art Portal, das einen zentralen Zugriff auf die Dienste ermöglicht.

⁶ Really Simple Syndication (RSS) [121] ist ein Format zur strukturierten Veröffentlichung von Änderungen auf Webseiten (z.B. Weblogs). Ein RSS-Feed stellt die Daten im RSS-Format bereit.

⁷ <http://www.elgg.org>, Zuletzt abgerufen am 29.01.2011

⁸ <http://www.flock.com>, Zuletzt abgerufen am 29.01.2011

⁹ <http://www.google.de/ig>, Zuletzt abgerufen am 29.01.2011

Wie gut eine PLE Lernprozesse und Wissensmanagement unterstützt, hängt von den eingebundenen Diensten und deren Konfiguration ab. Nur die Bereitstellung der Dienste ermöglicht noch kein effizientes Lernen und Wissensmanagement. Der Aufwand zum Einrichten und Konfigurieren einer PLE ist hoch und übersteigt eventuell die Kompetenzen mancher Lernender [7].

6.4 Social Bookmarking

Die Hauptfunktion von Social Bookmarking-Diensten ist die persönliche Organisation von Ressourcen wie z.B. Webseiten, Bildern oder Musik in Form von Bookmarks mittels Tagging (siehe Kapitel 3.1). Eine weitere Funktion ist das Weitergeben der Bookmarks, daher der Begriff "Social" im Namen. Anhand von Delicious¹⁰ wird in diesem Abschnitt die Funktionalität von Social Bookmarking-Diensten beschrieben. Delicious ist weit verbreitet und hat ähnliche Eigenschaften und Funktionalitäten wie die meisten anderen Social Bookmarking-Dienste. Es dient zum Speichern von Bookmarks zu Webseiten.

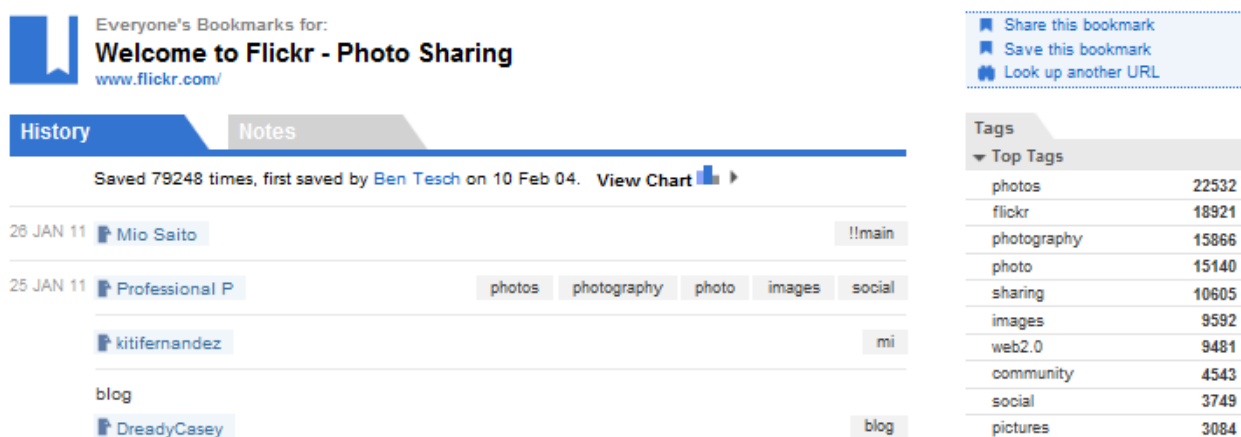


Abbildung 6.4: Bookmark zur Webseite von Flickr in Delicious

Die Bookmarks werden auf einem Server gespeichert, damit sind sie nicht mehr an den Web Browser des Nutzers gebunden. In Delicious werden die Bookmarks durch ihre URL identifiziert. Dabei kann jede URL von einem einzelnen Benutzer nur einmal abgespeichert werden, verschiedene Benutzer können jedoch die URL ebenfalls speichern. Abbildung 6.4 zeigt einen Ausschnitt für eine gespeicherte URL zur Webseite von Flickr. Der Eintrag besteht aus der URL und einem Titel für die Webseite. Außerdem wird eine chronologisch sortierte Liste aller Benutzer gezeigt, die die URL in Delicious gespeichert haben, und welche Tags diese Benutzer vergeben haben. Am rechten Rand werden Tags angezeigt, die besonders häufig für diese Webseite vergeben wurden. Die Zahl neben dem Tag gibt jeweils die Häufigkeit an. Zusätzlich können die Notizen der anderen Benutzer eingeblendet werden.

Abbildung 6.5 zeigt eine *Tag Cloud* der aktuell¹¹ häufigsten Tags aller Benutzer in Delicious. Tag Clouds sind Ansammlungen von Tags. Sie zeigen die Tags abhängig von der Häufigkeit ihrer Verwendung unterschiedlich an. Je häufiger das Tag vorkommt, desto größer wird das Wort angezeigt. Andere Varianten variieren die Farbe. Bei Delicious ist neben der alphabetischen Sortierung auch eine Sortierung der Tags anhand der Häufigkeit möglich.

¹⁰ <http://www.delicious.com>, Zuletzt abgerufen am 24.01.2011

¹¹ Stand: 26.01.2011

Tag Cloud: Popular

KEY: **green tags** are tags you have in common with everyone else.

Sort: Alphabetically | [By size](#)

.net 2008 3d advertising ajax and animation **api** apple architecture **art** article articles
artist audio **blog** blogging **blogs** book **books** browser **business** car **cms** code
collaboration comics community computer converter cooking cool **CSS** culture data
database **design** Design desktop **development** diy documentation download
downloads **drupal** ebooks economics **education** electronics email entertainment
environment fashion fic film finance **firefox** **flash** flex flickr **food** forum **free**
freeware fun funny gallery game **games** geek **google** government **graphics** green
guide hardware health history home hosting house **howto** **html** humor **icons**
illustration images imported information **inspiration** interactive interesting internet iphone
japan **java** **javascript** jobs jquery kids language learning library **linux** list lists
literature **mac** magazine management maps marketing **math** media microsoft mobile
money movie movies mp3 **music** network networking **news** online **opensource** osx
people phone photo **photography** photos photoshop php plugin podcast **politics**
portfolio privacy productivity **programming** psychology python radio rails realestate
recipe recipes **reference** religion **research** resources reviews **rss** ruby rubyonrails
school science **search** security seo shop **shopping** social socialnetworking
software **statistics** streaming teaching tech technology tips todo tool **tools** **toread**
travel tutorial tutorials tv twitter typography ubuntu **usability** **video** videos vim
visualization **web** **web2.0** **webdesign** webdev **wiki** wikipedia windows wishlist
wordpress work writing youtube

Abbildung 6.5: Tag Cloud in Delicious

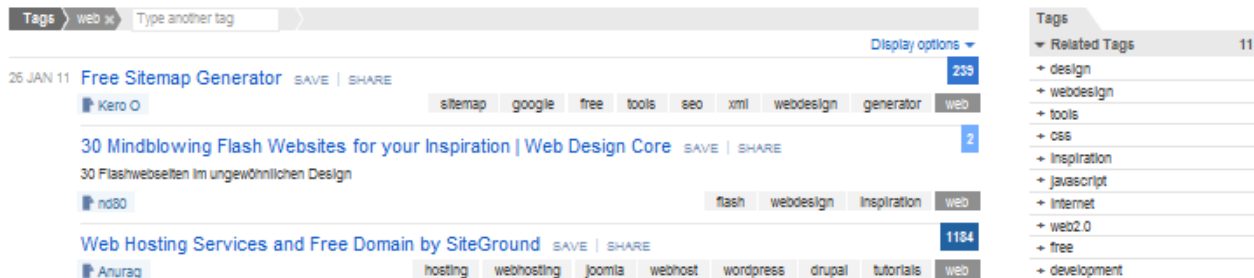


Abbildung 6.6: Browsen in Delicious mittels verwandten Tags

Tags dienen dazu, durch die eigenen Bookmarks und die anderer Benutzer in Delicious zu browsen (siehe Abbildung 6.6). Wird ein Tag ausgewählt, werden alle Webseiten aufgelistet, die mit diesem Tag versehen wurden. Die Zahlen in den Kästchen geben die Anzahl der Benutzer an, die die Webseite getaggt haben. Am rechten Rand werden Tags angezeigt, die besonders häufig gleichzeitig mit dem ausgewählten Tag zur Bezeichnung einer Webseite genutzt werden. Bei Auswahl eines dieser Tags werden alle Webseiten aufgelistet, die mit diesem Tag und dem zuerst gewählten Tag ausgezeichnet wurden. Jeder zusätzliche Tag verfeinert die Filterung. Des Weiteren können Tags nach ihrer Häufigkeit sortiert und solche ausgeblendet werden, die seltener als die ausgewählte Mindesthäufigkeit vorkommen. Sortier- und Filterfunktionen grenzen die Ergebnismenge ein.

Die Hierarchie der Tags in Delicious ist flach. Es gibt nur die Möglichkeit, Tags zu Gruppen, sogenannte Bundles, zusammenzufassen.

The screenshot shows the Delicious bookmarking form. At the top, there is a 'URL' field with the value 'http://getgreenshot.org/de/' and a 'Mark as Private' checkbox. Below this is a 'TITLE' field containing 'Greenshot - ein kostenloses open-source Screenshot-Programm für maximale Produktivität'. A 'NOTES' field is empty, with a '1000 chars' limit indicator. The 'TAGS' field is empty, with a help icon (?) to its right. Below the tags field is a 'SEND' field with the text 'Send your bookmarks' and a 'Clear' button. The 'Tags' section is expanded, showing two tabs: 'Tags' (selected) and 'Send'. Under the 'Tags' tab, there are two sections: 'Recommended Tags: click to add from your existing tags' with the tags 'screenshot' and 'tools', and 'Popular Tags: click to add from popular tags on Delicious' with the tags 'screenshot', 'opensource', 'tools', 'screenshots', and 'software'. At the bottom right, there are 'Save' and 'Cancel' buttons.

Abbildung 6.7: Taggen einer Webseite in Delicious

Zum Speichern der Bookmarks gibt es ein Add-on für den Webbrowser Mozilla Firefox. Wenn mit diesem Add-on eine neue Webseite zu Delicious hinzugefügt wird, öffnet sich das in Abbildung 6.7 gezeigte Eingabefenster. Automatisch werden der Titel der Webseite und die URL übernommen. Zusätzlich kann auf der Webseite Text markiert werden, der dann als Beschreibung übernommen wird oder es kann manuell eine Beschreibung eingegeben werden. Für das Bookmark gibt es zwei Sichtbarkeitsvarianten, entweder es ist nur für den angemeldeten Benutzer oder für alle Besucher von Delicious sichtbar. Im Eingabefeld für Tags können Schlagworte eingegeben werden. Das Trennzeichen für die Schlagworte ist das Leerzeichen, d.h. jedes Tag in Delicious kann nur aus einem Wort bestehen. Wollen Benutzer beispielsweise das Tag "Social Bookmarking" nutzen, dann müssen sie sich aushelfen, indem sie beispielsweise "SocialBookmarking" in CamelCase-Schreibweise eingeben (d.h. die Wörter werden zusammengezogen und der Anfangsbuchstabe der einzelnen Wörter groß geschrieben) oder das Leerzeichen durch andere Trennzeichen wie "-", "_", "." ersetzen (ergibt Social_Bookmarking). Diese Schreibweisen erschweren spätere Suchen.

Während der Eingabe der Tags werden die Benutzer durch eine Autovervollständigung auf Basis bereits existierender Tags unterstützt. Des Weiteren werden Tags vorgeschlagen (siehe Abbildung 6.7 unten), die von anderen Benutzern häufig für die zu speichernde Webseite vergeben wurden. Dabei wird unterschieden, ob das Tag vom angemeldeten Benutzer bereits für andere Webseiten genutzt wurde oder nur von anderen Benutzern.

In Delicious ist es möglich, ein soziales Netzwerk aufzubauen. Dazu wählt der Benutzer einen anderen Benutzer aus und fügt ihn so seinem Netzwerk hinzu. Für den anderen Benutzer erscheint dieser dann als Fan. Der Benutzer kann daraufhin verfolgen, welche Webseiten von den Mitgliedern seines Netzwerks gespeichert werden. Außerdem kann der Benutzer den Mitgliedern in seinem Netzwerk direkt Webseiten mit einer kurzen Notiz empfehlen. Die Empfehlungen tauchen dann im Posteingang ("Inbox") auf. Neben der Weitergabe direkt in Delicious gibt es auch die Möglichkeit, die Empfehlung per Mail oder Twitter weiterzugeben. Des Weiteren ist es möglich, Einträge zu abonnieren, d.h. bestimmte Tags zu ver-

folgen. Das Verfolgen von Einträgen kann auch per RSS-Feed geschehen. Zusätzlich stehen Entwicklern Programmierschnittstellen zur Verfügung, um die Einträge extern verfügbar zu machen.

Diskussion

Social Bookmarking eignet sich gut für die Organisation von Web-Ressourcen; genau für diesen Einsatzzweck wurde es konzipiert. Der Mehraufwand zum Weitergeben der Bookmarks ist gering, er entspricht in etwa dem Hinzufügen eines Tags, da der Name des Empfängers auf die gleiche Art und Weise wie ein Tag hinzugefügt wird. Das Finden von neuen Ressourcen wird durch das Filtern nach Tags und durch Abonnements unterstützt, dabei können aber nur Webseiten gefunden werden, die von anderen Benutzern bereits in Delicious gespeichert wurden. Ein Problem in Delicious stellen die 1-Wort-Tags dar, da sie die Schlagworte entweder auf ein einzelnes Wort begrenzen oder zu alternativen Schreibweisen führen. Delicious wurde nicht explizit für die Unterstützung von Lernszenarien entworfen. Die Planung und Reflektion des Lernprozesses ist nicht vorgesehen, wenngleich eine Art Aufgabenmanagement durch Tags realisiert werden kann. Dies kann beispielsweise durch Tags geschehen, die Aktionen beschreiben wie "to_read" oder die das Projekt beschreiben, in dessen Rahmen die Webseite gefunden wurde. Die Möglichkeiten zur Annotation der Webseiten sind auf Tags und einfache Textnotizen mit einer maximalen Länge von 1000 Zeichen beschränkt. Die Webseite wird nur als Link gespeichert. Eine Volltextsuche in den verlinkten Webseiten ist damit nicht möglich. Somit können bei einer Suche nur Informationen gefunden werden, die die Benutzer in Delicious gespeichert haben. Eine Ordnung der Tags ist nur durch Bundles möglich. Die flache Hierarchie trägt kaum zur Wissenskonstruktion bei. Die Nutzung der Informationen ist auf Suche, Darstellung und einfache Exportfunktionalität begrenzt.

6.5 Zielmanagement

Neben einfachen Applikationen wie Kalender und To-do-Listen gibt es Ansätze, die die Lernenden bei der Planung und Reflektion des Lernprozesses auf Basis der Selbstregulationstheorie unterstützen. Selbstregulation kann definiert werden als *selbst erzeugte Gedanken, Gefühle und Verhaltensweisen, die basierend auf dem Feedback über die Leistung bei der Erreichung selbst-gesetzter Ziele geplant und zyklisch angepasst werden* [166]. Im Allgemeinen werden drei zyklische Phasen genannt [32]: Planung (Forethought), Ausführung (Performance) und Selbstreflektion (Self-Reflection). In der Planungs-Phase findet das Setzen von Zielen statt und das Festlegen der Strategie zur Erreichung des Zieles. In der Ausführungs-Phase, in der gelernt wird, beobachten die Lernenden ihren Lernprozess und Lernfortschritt. In der Selbstreflektions-Phase werten die Lernenden ihren Lernerfolg und ihre Strategien aus.

Beim selbst-gesteuerten Lernen ist es wichtig, dass die Lernenden für ihren Lernprozess Verantwortung übernehmen. In [135] wird eine Applikation für das Lernen von Englisch vorgestellt. Die Lernumgebung besteht aus einem Buch, das Englischlektionen enthält, und einem Hyperpen. Der Hyperpen ist ein Stift, mit dem Textstellen gescannt und zur Software übertragen werden können. Die Software ermöglicht Zugriff auf Wörterbücher und weitere Lernmaterialien. In der Software können sich die Lernenden ihre Lernziele setzen und einen Zeitplan zur Erreichung der Lernziele erstellen. Der Zeitplan wird als wöchentlicher Kalender dargestellt, in den die Lernenden geplante Aktivitäten eintragen können (was, wann und wie lange). Mit diesem Kalender können sie somit ihre Strategie planen. Während sie lernen,

können sie Aktivitäten auswählen, die sie bearbeiten wollen. Daraufhin erhalten sie die zu den Aktivitäten zugeordneten Lernmaterialien. Während der Bearbeitung können sie Unterbrechungen eintragen. Das Reflektieren findet mittels statistischer Auswertungen statt. Die Lernenden können sich Diagramme anzeigen lassen über die Häufigkeit des Lernens und die Abweichungen gegenüber ihrem Zeitplan. Des Weiteren können sie in der Software Selbsttests zur Messung ihres Lernfortschritts durchführen. Die Software ermöglicht es, dass ein Lehrer Beispiel-Zeitpläne für die Lektionen erstellt, die die Schüler herunterladen und anpassen können.

Eine Applikation, die auch auf vordefinierte Lektionen festgelegt ist, ist gStudy [162]. In dieser Applikation können die Lernenden Notizen, Glossare und Concept-Maps erstellen, Inhalte annotieren, nach Informationen suchen, chatten und kollaborieren und Hilfe anfordern. Im Gegensatz zur o.g. Lernumgebung zeichnet gStudy alle Aktivitäten mit der Software auf. Die aufgezeichneten Daten sollen die Lernenden bei der Reflektion unterstützen, beispielsweise durch die Anzeige der Häufigkeit der Durchführung einer Aktivität.

Des Weiteren existieren Ansätze, die auf die Unterstützung kollaborativen Lernens fokussieren und die den Austausch zwischen Lernenden unterstützen wollen. In [137] wird eine Social Bookmarking-Applikation vorgestellt, die den Austausch von Informationen in Organisationen ermöglichen soll. Benutzer können Ziele anlegen, und zu diesen Zielen dann Webseiten sammeln. Zusätzlich können die Benutzer die Ziele und Webseiten auch mit anderen teilen. Es gibt unterschiedliche Gründe für Benutzer, ihr Wissen mit der Organisation zu teilen, aber es gibt auch Hemmungen, dies zu tun. Manche Menschen haben Angst, dass sie ihren Arbeitsplatz verlieren, wenn sie all ihre Erfahrungen dokumentieren und mit anderen teilen, da sie dann ersetzt werden könnten. Der Fokus der Arbeit in [137] liegt auf der Analyse von Motivation und Ablehnung von Nutzern, ihr Wissen mit anderen zu teilen, um eine entsprechende Software-Lösung zu gestalten. Beispielsweise können Mechanismen die Motivation erhöhen, die den Wissensaustausch eines Nutzers sichtbar machen und damit seine Anerkennung in der Organisation erhöhen. Eine ähnliche Arbeit [84] fokussiert auf den Austausch von Zielen mit anderen Nutzern, um Verbindungen zwischen Nutzern mit ähnlichen Lernzielen herzustellen und um gemeinsam Webseiten für Ziele sammeln zu können.

Diskussion

Die Lernumgebungen in diesem Abschnitt fokussieren auf die Unterstützung des Zielmanagements beim Lernen. Einige der Ansätze integrieren wesentliche Funktionen zur Selbstregulation. Dazu gehören das Setzen von Zielen, die Planung der Lernstrategie, die Beobachtung der Lernenden während der Durchführung der Lernaktivitäten und die Reflektion des Lernprozesses bzw. -ergebnisses. Als Hilfsmittel in den Ansätzen dienen Kalenderansichten und Auswertungsdiagramme, die entweder manuell erstellt oder automatisch aufgezeichnet, die Aktivitäten der Lernenden widerspiegeln. Hilfe können die Lernenden dabei von Lehrern oder der Lernsoftware erhalten oder durch den Austausch mit anderen Nutzern, die ähnliche Lernziele verfolgen.

6.6 Fazit

In diesem Abschnitt wurden verwandte Arbeiten aus fünf verschiedenen Bereichen vorgestellt. Neben diesen gibt es weitere Ansätze, die Teile des Ressourcenmanagements unterstützen, beispielsweise Software zur Verwaltung von Literatur oder E-Portfolios. Die Ansätze bringen aber keine weiteren Vorteile gegenüber den hier bereits vorgestellten Ansätzen oder sind teilweise bereits in die hier vorgestellten Ansätze integriert, so dass an dieser Stelle auf eine eingehende Beschreibung verzichtet wird.

Im Folgenden werden die Ansätze hinsichtlich der Anforderungen der Aktivitäten des Ressourcenmanagements im Anwendungsszenario systematisch verglichen. Dabei wird von den Basisfunktionen, wie sie zuvor vorgestellt wurden, ausgegangen. Oftmals erfolgen in konkreten Anwendungen Kombinationen der Ansätze, was hier nicht berücksichtigt wird.

- **Planen & Reflektieren:** Semantic Desktops bieten zwar die technische Grundlage, um eine Applikation zur Planung und Reflektion zu realisieren, aber die Applikationen gehen bisher nicht über Kalender und To-do-Listen hinaus. In Social Bookmarking-Applikationen können zwar Tags vergeben werden, die dem Zielmanagement dienen, aber deren Funktionalität in Hinblick auf Zielmanagement ist stark eingeschränkt. Conceptual Diagrams können dabei helfen, den Wissensstand des Lernenden zu visualisieren und Wissenslücken aufzudecken (Reflektion), aber sie enthalten keine Planungsfunktionen. In den Zielmanagement-Ansätzen wird das Planen und Reflektieren des Lernprozesses in größerem Umfang unterstützt.
- **Suchen:** Keiner der Ansätze unterstützt ausreichend die Suche nach neuen Ressourcen im Web. In Semantic Desktops gibt es semantische Suchen, die das Durchsuchen der eigenen Informationssammlung unterstützen. In Social Bookmarking-Applikationen kann mittels Tags gefiltert werden.
- **Annotieren & Organisieren:** Jeder der Bereiche unterstützt Funktionen zur Annotierung und Organisation von Informationen, fokussiert aber auf unterschiedliche Aspekte. Conceptual Diagrams fokussieren auf die Anordnung und Visualisierung, Semantic Desktops auf die Erstellung eines persönlichen Wissensmodells, PLEs auf die Zusammenstellung einer personalisierten Lernplattform, Zielmanagementsoftware auf die Organisation mittels Zielen und Social Bookmarking auf Tagging.
- **Nutzen:** Zielmanagement-Software und Social Bookmarking-Dienste haben keine Funktionen in Bezug auf diesen Prozessschritt, die über eine Nutzung von Informationen nach einer Suche in der Applikation hinausgehen. Dienste zur Erstellung von Inhalten sind in PLEs auf Basis der Schnittstellen eng verbunden mit Diensten zur Organisation der Inhalte.
- **Weitergeben:** Conceptual Diagramms können zwar weitergegeben werden, aber sie sind ohne Erklärungen nicht immer verständlich für andere. Der Fokus der anderen Ansätze liegt entweder auf der aktiven Weitergabe durch die Nutzer oder auf der automatischen Weitergabe durch Entwicklung eines allgemeinen Verständnisses.

Jeder der fünf vorgestellten Ansätze unterstützt bestimmte Prozessschritte im Anwendungsszenario des Ressourcen-basierten Lernens, keiner erfüllt aber alle Anforderungen. Für eine Unterstützung des Ressourcen-basierten Lernens wird in dieser Arbeit eine Kombination und Anpassung der hier vorgestellten Konzepte zu einer persönlichen bzw. Community-Lernumgebung vorgeschlagen. In Tabelle 6.1 markiert ein "X" die Stärke des Ansatzes, die in die Kombination einfließt. Das Konzept dieser Lernumgebung wird in Kapitel 7 detailliert vorgestellt.

	Conceptual Diagram	Semantic Desktop	Personal Learning Environment	Zielmanagement- Software	Social Bookmarking
Planen & Reflektieren	(X)			X	
Suchen		Semantische Suche			Tags
Annotieren & Organisieren		X			X
Nutzen			X		
Weitergeben					X

Tabelle 6.1: Vergleich der verwandten Ansätze hinsichtlich der Anforderungen des Ressourcenmanagements im Ressourcen-basierten Lernen



7 Basiskonzept und Design des Unterstützungswerkzeugs

Dieses Kapitel beschreibt das im Rahmen der Arbeit entwickelte Konzept für das Unterstützungswerkzeug. Ein Werkzeug für das Ressourcen-basierte Lernen muss alle identifizierten Prozessschritte in dem im Kapitel 2.3.3 beschriebenen Modell unterstützen. Dazu muss es verschiedene Funktionen realisieren, diese werden im Abschnitt 7.1 erläutert. Des Weiteren werden bestimmte technische Anforderungen an ein solches Werkzeug gestellt, diese werden in Abschnitt 7.2 erarbeitet. Im Fokus steht dabei die Unterstützung des individuellen Lernens. Für das Lernen in Gruppen werden Erweiterungen des Konzepts dennoch kurz vorgestellt.

7.1 Funktionale Anforderungen und Konzept

Im Modell für das Ressourcen-basierte Lernen wurden fünf wichtige Prozessschritte identifiziert. Funktionen zur Unterstützung dieser Prozessschritte werden in diesem Abschnitt erläutert.

Planen & Reflektieren

Das Ressourcen-basierte Lernen erfordert von selbstgesteuert Lernenden ein hohes Maß an Selbstorganisation und -motivation. Schmitz und Wiese [127] gliedern einen selbstgesteuerten Lernvorgang in drei Phasen:

- den Zeitraum vor dem Lernen,
- den Zeitraum während des Lernens und
- den Zeitraum nach dem Lernen.

Des Weiteren beschreiben sie verschiedene Prozesse, die in diesen Zeiträumen stattfinden sollten. Beispielsweise sollten sich Lernende vor dem Lernen Ziele setzen und ihr Vorgehen planen. Während des Lernens sollten die Lernenden ihr Vorgehen beobachten und es an den Verlauf des Lernvorganges anpassen. Im Zeitraum nach dem Lernen sollten Lernende über ihr Vorgehen reflektieren und prüfen, ob sie ihr Vorgehen möglicherweise verbessern können. Diese Prozesse lassen sich auf die Recherche im Web und das Ressourcen-basierte Lernen übertragen.

Ein Werkzeug soll demnach Funktionen anbieten, die diese Prozesse ermöglichen und anregen. Das Konzept für das Unterstützungswerkzeug sieht daher folgende Funktionen vor:

- **Ziele setzen:** Für die Planung des Vorgehens soll eine Funktion zur Verfügung stehen, mit der sich die Lernenden Ziele setzen können. Jedes Ziel muss durch einen Bezeichner identifizierbar sein, hat einen Erstellungszeitpunkt und einen erstellenden Nutzer. In einer Beschreibung können Lernende Notizen zum Ziel festhalten. Alternativ könnte das Werkzeug auch spezielle Eingaben fordern, beispielsweise die SMART-Kriterien eines Ziels. SMART ist eine Abkürzung für "Spezifisch", "Messbar", "Anspruchsvoll", "Realistisch" und "Terminierbar" [79]. Diese Ziel-Kriterien können bei der Formulierung von Zielen helfen. Spezifisch bedeutet, dass Ziele eindeutig definiert sein sollen,

nicht vage. Nur bei messbaren Zielen kann man deren Erfolg auch nachvollziehen. Bei Zielen, die nicht anspruchsvoll genug oder zu anspruchsvoll sind, sinkt die Motivation, daher sollten Ziele angemessen formuliert werden. Ziele sollten realistisch erreichbar sein und sie sollten ein klares Ende haben, d.h. einen festen Termin zur Fertigstellung. Ein Ziel im Anwendungsszenario könnte folgendermaßen lauten: "Ich will bis zum 1. April zehn Ressourcen zum Thema XY gesammelt haben."

- **Ziele hierarchisch strukturieren:** Ziele wie das Schreiben einer Hausarbeit oder die Erstellung einer wichtigen Folienpräsentation umfassen in der Regel eine längere Zeitspanne, daher ist es wichtig, dass die Ziele hierarchisch in Teilziele unterteilt werden können.
- **Ressourcen zu Zielen zuordnen:** Beim Ressourcen-basierten Lernen wird es Ziele geben, für die neue Informationen beschafft werden müssen. Daher werden Lernende für diese Ziele Ressourcen sammeln, was zentrale Aufgabe des Werkzeugs ist. Die Ressourcen müssen mit dem entsprechenden Ziel verbunden werden können.
- **Ziele umstrukturieren:** Während des Lernens können sich Änderungen ergeben, die sich auf die geplanten Ziele auswirken können. Deshalb muss das Werkzeug eine Umstrukturierung jederzeit erlauben, d.h. es muss möglich sein, Ziele umzubenennen oder untergeordnete Ziele anderen übergeordneten Zielen zuzuordnen. Dabei müssen die einem Ziel zugeordneten Ressourcen unverändert bleiben.
- **Aktiv Ziele verfolgen:** Während der Bearbeitung der Lernaufgabe sollte der Lernende konzentriert ein Ziel verfolgen und sich nicht von anderen noch zu erledigenden Aufgaben oder momentan irrelevanten Webseiten ablenken lassen. Eine Funktion, die dem Lernenden sein aktuelles Ziel bewusst macht, hilft dem Lernenden Ablenkungen zu minimieren und strukturierter vorzugehen.
- **Überblick über den Fortschritt der Zielbearbeitung gewinnen:** Es muss außerdem Funktionen geben, mit der sich die Lernenden einen Überblick über ihre bereits vorhandenen Informationen zu einem Ziel verschaffen können, um Informationslücken identifizieren zu können. Informationen sind im Ressourcen-basierten Lernen insbesondere die einem Ziel zugeordneten Ressourcen. So muss beispielsweise übersichtlich dargestellt werden, wie viele und welche Ressourcen einem Ziel zugeordnet sind. Ein Überblick über den Fortschritt eines Zieles ist insofern auch sinnvoll, weil die Bearbeitung der Lernaufgabe manchmal zeitlich unterbrochen werden muss. Bei Wiederaufnahme der Bearbeitung kann der Lernende dann Anknüpfungspunkte an die bereits begonnene Arbeit finden.
- **Reflektieren über das Vorgehen:** Das Anzeigen von Ressourcen, die während einer bestimmten Zeit gesammelt wurden, kann den Lernenden zur Reflektion seines bisherigen Vorgehens anregen. Denn werden beispielsweise keine Ressourcen angezeigt, bedeutet dies für den Lernenden, dass er seine Suchstrategie verändern muss, um sein Ziel erfolgreich fertigzustellen.
- **Status der Zielerreichung angeben:** Nach dem Lernen bzw. bei Erreichen von Teilzielen kann ein Abhaken des Zieles motivierend sein. Zudem erhöht das die Übersichtlichkeit über den Fortschritt der Zielbearbeitung. Der Lernende muss also angeben können, ob er das Ziel erreicht hat bzw. zu welchem Grade er das Ziel erreicht hat.
- **Bearbeitungszeiträume planen:** In den meisten Fällen hat ein Lernender verschiedene Ziele und gegebenenfalls noch andere feste Termine wie z.B. Dienstreisen oder Meetings. Ein Kalender kann bei der Planung von Zielen helfen, indem freie Zeiträume sichtbar werden. In diesen Freiräumen

können Termine für die Bearbeitung der Ziele geblockt werden. Das Werkzeug kann hierbei eine Anbindung an einen Kalender ermöglichen, indem es Kalendertermine automatisch als Ziel übernimmt oder in der umgekehrten Richtung Ziele als Kalendereinträge exportiert.

Community-Funktionen

Beim Lernen in Lerngruppen gibt es Ziele, die gemeinsam erreicht werden sollen. Folgende Funktionen, die Lerngruppen unterstützen, können das Konzept für ein Unterstützungswerkzeug erweitern:

- Andere Lernende kontaktieren, d.h. beispielsweise Nachrichten austauschen und Ziele kommentieren,
- Beobachten von Aktionen anderer Lernender der Lerngruppe, beispielsweise welche Ziele haben andere Lernende sich gesetzt, welche Ressourcen wurden von ihnen gesammelt und wie wurden diese Ressourcen getaggt,
- Erstellen gemeinsamer Ziele, zu denen gemeinsam Informationen recherchiert und aufgearbeitet werden und
- Aufteilen der Ziele auf verschiedene Mitglieder der Lerngruppe, so dass eine Verteilung der Aufgabe stattfinden kann.

Annotieren & Organisieren

Eine der zentralen Aufgaben des Werkzeugs ist die Verwaltung der Ressourcen im Wissensnetz. Daher sieht das Konzept Funktionen vor, die das Speichern der Webseiten als Ressourcen und das Organisieren durch Tagging wie folgt ermöglichen:

- **Webseiten speichern:** Bei der Recherche nach neuen Informationen im Web soll es möglich sein, Webseiten direkt im Wissensnetz abzuspeichern. Der Lernende soll allerdings nicht gezwungen werden, eine Vielzahl verschiedener Metadaten auszufüllen. Deshalb soll vom Werkzeug nur eine geringe Anzahl von Metadaten verpflichtend vorgegeben sein, die der Lernende individuell erweitern kann. Wichtige Metadatenattribute für Ressourcen im Wissensnetz hierbei sind:
 - **URL:** Die URL der Webseite sollte automatisch vom Werkzeug erfasst werden.
 - **Bezeichner (Titel) der Ressource:** Der Titel der Webseite kann vom Werkzeug automatisch erfasst werden und als Bezeichner der Ressource im Wissensnetz dienen, soll aber vom Lernenden änderbar sein.
 - **Zeitpunkt des Speicherns:** Der Zeitpunkt des Speicherns der Webseite im Wissensnetz soll vom Werkzeug automatisch erfasst werden.
 - **Snippets:** Teilweise sind Auszüge aus der zu speichernden Webseite für den Lernenden besonders relevant. Die Auszüge werden in dieser Arbeit als "Snippet" bezeichnet. Diese Auszüge sollen durch den Lernenden auswählbar sein und als Metadatum der Ressource im Wissensnetz gespeichert werden. Das Anpassen des Snippets durch den Lernenden soll hierbei erlaubt sein. Des Weiteren kann der gesamte Inhalt der Webseite in einem Cache gespeichert werden, um auch bei Verschwinden der Webseite z.B. durch Löschen der Webseite oder durch zeitweisen Ausfalls des Servers, auf dem die Webseite liegt, den Zugriff auf die Inhalte sicherzustellen.

- **Screenshot:** Das Werkzeug kann automatisch einen Screenshot der zu speichernden Webseite erstellen. Die Screenshots können dann als Vorschau dienen und beim Wiederfinden der Ressource als Erinnerungstütze helfen.
- **Annotieren:** Bei Bedarf soll ein Administrator oder auch jeder Lernende individuell die Metadatenattribute erweitern können. Dazu soll das Werkzeug es erlauben, weitere Metadaten zu definieren. Beispielsweise ist der Fall denkbar, dass ein Lernender zusätzlich zu den bestehenden Metadaten noch eine Zusammenfassung oder Angaben zur Qualität der Webseite oder zur Relevanz für die Lösung der Aufgabe speichern will.
- **Typisiert taggen:** Neben der Beschreibung mit Metadaten soll das Werkzeug auch das typisierte Tagging (siehe Kapitel 4) unterstützen. Tags sollen frei vom Lernenden vergeben und beim Taggen einem der Tag-Typen zugeordnet werden können. Auch eine nachträgliche Zuordnung zu einem Tag-Typ soll möglich sein. Es soll ebenso vorgesehen werden, dass Tags annotiert werden können z.B. durch Beschreibungen, und auch Relationen zwischen Tags angelegt werden können, beispielsweise Unter-/Oberrelationen zwischen Themen.
- **Tagvorschläge:** Bei der Eingabe eines Tags soll es eine Autovervollständigen-Funktion geben, d.h. nach Tippen der ersten Buchstaben wird eine Liste mit passenden, bereits im Wissensnetz gespeicherten Tags eingeblendet. Des Weiteren kann das Werkzeug beim Taggen der Ressource Vorschläge für Tags machen (siehe Kapitel 4.4), die mit wenig Aufwand vom Lernenden übernommen werden können. Falls für diese Tags Zusatzinformationen existieren, sollen diese bei Bedarf eingeblendet werden.
- **Wissensnetz pflegen:** Beim Speichern der Ressourcen macht sich der Lernende Gedanken, wie er die Ressource im Wissensnetz abspeichern will, d.h. mit welchen Schlagworten er die Ressource taggt. Dabei hat er vor allem die aktuelle Recherche im Blick und weniger das ganze Wissensnetz. Daher kann es vorkommen, dass teilweise nicht strukturiert getaggt wird und dass Bedarf für Pflege und Umstrukturierung des Wissensnetzes besteht. Mit folgenden Funktionen soll das Werkzeug die Pflege des Wissensnetzes unterstützen:
 - **Bearbeitung im Graph:** Es soll nicht nur das Wissensnetz als Graph dargestellt werden, sondern auch die Bearbeitung des Wissensnetzes in der graphischen Darstellung ermöglicht werden. Dazu gehört neben elementaren Bearbeitungen wie dem Umbenennen von Tags oder dem Erstellen neuer Tags beispielsweise auch die Möglichkeit, mehrere Ressourcen im Wissensnetz auszuwählen und gleichzeitig mit einem Tag zu verschlagworten.
 - **Zusammenfassen von ähnlichen Tags:** Tags, die zusammengefasst werden können, können vorgeschlagen werden, beispielsweise Synonyme, Singular- und Pluraltags, Tags in verschiedenen Sprachen, Abkürzungen etc.
 - **Themenstruktur aufbauen:** Ein Lernender, der sich ein Themengebiet neu aneignet, kennt eventuell nur sehr generelle Konzepte des Themas. Während der Recherche lernt er dann das Themengebiet näher kennen und dadurch auch speziellere Konzepte, die die Ressource genauer beschreiben. Je größer die Menge an Ressourcen ist, die mit einem Thema getaggt wurden, desto schwieriger ist oftmals das Wiederfinden. Der Lernende soll die Möglichkeit haben, die Themen weiter zu untergliedern und die Menge der Ressourcen mit diesen spezielleren Themen zu verschlagworten. Verfahren wie XESA [131] können anhand der Ähnlichkeiten

der Ressourcen mittels externer Datenquellen Themenvorschläge generieren, indem sie z.B. den Ressourcen ähnliche Wikipedia-Artikel identifizieren und dessen Wikipedia-Kategorien als neue Themen vorschlagen.

- **Ergänzung der Tags durch automatisch extrahierte Schlagworte:** Manche der Tags können automatisch bestimmt werden, beispielsweise gibt es Verfahren zur Webgenre-Erkennung [130], die dem Lernenden automatisch Vorschläge für den Tag-Typ "Typ" machen können, indem sie das Webgenre der Webseite identifizieren, z.B. ob die Webseite ein Weblog, Wiki oder Forum ist.
- **Relationen zwischen Tags:** Des Weiteren kann das Werkzeug Relationen zwischen Tags vorschlagen. Wie in Kapitel 5.2.2 erläutert, enthalten manche Datenquellen Konzepte, die untereinander verlinkt sind. Diese Links können als Relationen vorgeschlagen werden.

Suchen

Informationen, die für die Bearbeitung eines Ziels notwendig sind, können bereits als Ressourcen im Wissensnetz gespeichert worden sein - und zwar vom Lernenden selbst oder von anderen Lernenden. Daher sollte das Werkzeug folgende Möglichkeiten bieten, um das Wissensnetz zu durchsuchen:

- **Volltextsuche :** Es soll eine Volltextsuche zur Verfügung stehen, die in allen textbasierten Inhalten des Wissensnetzes nach den vom Lernenden eingegebenen Suchbegriffen sucht. Diese Volltextsuche soll sich bei Bedarf auf bestimmte Attribute einschränken lassen, beispielsweise um nur nach den Bezeichnern von Ressourcen zu suchen.
- **Typisierte Suche:** Das Werkzeug soll eine strukturelle Suche anhand der typisierten Tags anbieten, damit Suchen wie "Finde alle Ressourcen die mit dem Ziel Z getaggt wurden" möglich sind. Auch erweiterte Suchen über Relationen zwischen Tags wie "Finde alle Ressourcen die mit Konferenzen zum Thema XY zusammenhängen" werden durch das Konzept des typisierten Taggings möglich.
- **Suchergebnisse darstellen:** Für die Darstellung der Suchergebnisse sind verschiedene Ansichten sinnvoll: tabellarische Anzeige, spezielle Ansichten abhängig vom Typ z.B. kalenderbasierte Anzeige für Ereignisse, Anzeige als Graph usw. Auf die verschiedenen Darstellungen des Wissensnetzes wird in Abschnitt "Visualisierung" näher eingegangen. Bei der Anzeige von Suchergebnissen sollen Inhalte, die vom Suchenden selbst im Wissensnetz gespeichert bzw. erstellt wurden, anders dargestellt werden als Inhalte von anderen Lernenden.

Für manche Ziele wird der Lernende neue Ressourcen im Web suchen. Die Suchmöglichkeiten sollen dabei nicht vom Werkzeug eingeschränkt werden, sondern durch zusätzliche Informationen unterstützt werden. Informationen, die das Konzept für die Suche im Web vorsieht, sind die folgenden:

- **Hinweise zu im Wissensnetz gespeicherten Ressourcen:** Wenn eine Webseite im Browser geöffnet wird, die bereits im Wissensnetz als Ressource gespeichert wurde, soll das Werkzeug darüber informieren, dass sie bereits gespeichert wurde und eventuell von welchen Lernenden. Diese bereits gespeicherten Ressourcen können wiederum Startpunkt für eine Navigation im Wissensnetz sein, daher soll dieser Hinweis auf gespeicherte Ressourcen auch einen Verweis auf die entsprechende Ressource im Wissensnetz enthalten.

-
- **Gespeicherte Snippets:** Des Weiteren kann es, wie in Abschnitt "Annotieren & Organisieren" erläutert, möglich sein, dass Lernende Teile der Webseite als Snippet abgespeichert haben. Diese Snippets sollen auf der entsprechenden Webseite hervorgehoben werden, beispielsweise durch farbliche Markierung der entsprechenden Abschnitte ähnlich zur Highlighting-Funktion der Webanwendung Diigo¹.
 - **Vorschlagen von weiteren Ressourcen:** Wurde eine Webseite bereits gespeichert und getaggt, ist es möglich, dass es weitere ähnliche Ressourcen im Wissensnetz gibt, d.h. beispielsweise Ressourcen zum gleichen Thema (siehe auch Kapitel 4.4). Diese verwandten Ressourcen sollen den Lernenden angezeigt werden, da sie möglicherweise für die aktuelle Suche relevant sind.
 - **Ranking von Suchergebnissen:** Möglich sind neben der Einblendung von Zusatzinformationen für die aufgerufenen Webseiten auch das Ranking oder Filtern von Suchtreffern. Beispielsweise kann die Webseite einer Suchmaschine, die die Suchtreffer auflistet, dahingehend manipuliert werden, dass abhängig vom Wissensnetz des Lernenden die Suchtreffer neu gerankt werden oder bestimmte Suchtreffer ausgeblendet werden. Sucht ein Lernender z.B. nach dem Begriff "Java" und hat im Wissensnetz das Ziel "Programmieren lernen" aktiviert, könnten auf dieser Basis Suchtreffer über die Insel Java ausgeblendet werden.

Nutzen

- **Verschiedene Ansichten:** Das Werkzeug soll das Wissensnetz in verschiedenen Ansichten darstellen. Mögliche Ansichten sind in Abschnitt "Visualisierung" aufgezählt.
- **Lesereihenfolge festlegen:** Nachdem eine Reihe neuer Ressourcen recherchiert wurde, die eventuell nur überflogen wurden, kann es für den Lernenden hilfreich sein, wenn er die Ressourcen in eine Art Lesereihenfolge bringt, die er dann nacheinander abarbeiten kann. Das Werkzeug soll dem Lernenden hierfür eine Funktion anbieten, mit der er eine Art "Pfad durch seine Ressourcen" erstellen kann.
- **Exportfunktionen:** Wenn Inhalte aus dem Wissensnetz genutzt werden, besteht häufig der Bedarf die Quelle der Informationen anzugeben, beispielsweise in Folienpräsentationen. Das Werkzeug soll daher geeignete Exportfunktionen anbieten wie eine strukturierte Liste ausgewählter Snippets, um diese in einer Textverarbeitungssoftware weiter zu verarbeiten; oder beispielsweise auch die Generierung eines Eintrages der URL im Literaturverzeichnis oder Fußnoten.

Weitergeben

Das Empfangen und Weitergeben von Inhalten aus dem Wissensnetz an bzw. von anderen Lernenden soll mit angemessenem Aufwand möglich sein:

- **Implizites Weitergeben:** Ohne explizites Aufbereiten der Inhalte des Wissensnetzes soll ein implizites Weitergeben möglich sein, d.h. ein Lernender soll sich die Wissensnetze anderer Lernender anschauen können, wenn jene Lernende den Lesezugriff erlaubt haben.

¹ <http://www.diigo.com/>, Zuletzt abgerufen am 28.02.2011

- **Export ausgewählter Inhalte:** Inhalte sollen auch an Personen, die keinen Zugriff auf das Wissensnetz haben, weitergegeben werden können. Dazu sollte das Werkzeug Möglichkeiten für den Export ausgewählter Inhalte haben, z.B. alle mit einem ausgewählten Ziel verknüpften Ressourcen als Linkliste exportieren.
- **Explizite Empfehlungen:** Das Weitergeben von Informationen an Personen, die Zugriff auf das Wissensnetz haben, soll durch Empfehlungen möglich sein. Ausgewählte Ressourcen oder Tags könnten mit einem Klick in eine Art Posteingang des Lernenden wandern, der die Empfehlung erhalten soll. In Delicious gibt es dafür das spezielle Tag "for:XY". Die getaggtten Ressourcen sind dann in einem speziellen Bereich des Delicious-Accounts zu finden. Auch der im Prozessschritt "Nutzen" erwähnte "Ressourcenpfad" soll weitergegeben werden können bzw. zum Zwecke der expliziten Weitergabe von Inhalten erstellt werden können.
- **Abonnieren:** Explizite Empfehlungen gehen von den Lernenden aus, die die Ressourcen im Wissensnetz speichern, aber es soll auch der umgekehrte Fall möglich sein, bei dem Lernende die Inhalte anderer Lernender abonnieren, um über Änderungen oder neue Inhalte informiert zu werden. Dabei soll es möglich sein, dass einzelne Ressourcen abonniert werden, für die der Lernende bei neuen Kommentare informiert werden will, oder dass Tags abonniert werden, für die der Lernende Informationen zu neu gespeicherten Ressourcen, die mit diesem Tag verschlagwortet wurden, erhalten will.
- **Mapping der persönlichen Wissensnetze:** Beim Betrachten der Wissensnetze anderer Lernender wird es manchmal Orientierungsprobleme geben, weil die Lernenden teilweise verschiedene Tag-Bezeichner nutzen werden. Automatische Mapping-Verfahren (siehe [136, 30, 43] für einen Überblick) versuchen semantische Strukturen aufeinander abzubilden (Synonyme, Homonyme etc.). Solche Verfahren könnten dabei helfen, die Wissensnetze anderer Lernender an die Struktur des Wissensnetzes des Lernenden, der die anderen Wissensnetze betrachtet, für die Darstellung anzugleichen. Beispielsweise ist das Zusammenfassen von sehr speziellen Themen zu einem Thema denkbar, das auch im Wissensnetz des Lernenden vorhanden ist, z.B. wenn der Lernende nur das Thema "Wind" in seinem Wissensnetz hat, aber in den fremden Wissensnetzen die Ressourcen mit "Orkan", "Sturm" und "Tornado" getaggt sind.

Community-Funktionen

- **Leserechte:** Bei der Weitergabe von Inhalten des Wissensnetzes ist es möglich, dass die Lernenden den Zugriff einschränken wollen und deshalb Leserechte festlegen möchten. Community-Funktionen wie das Erstellen von Freundeslisten und das Bilden von Lerngruppen können Leserechte für spezielle Gruppen von Lernenden realisieren.
- **Vorschlagen von Lerngruppen:** Auf der Basis von ähnlichen Themen oder ähnlichen Ressourcen könnte das Werkzeug dem Lernenden andere Lernende für die Bildung von Lerngruppen vorschlagen, um Aktionen wie im Abschnitt der Community-Funktionen im Prozessschritt "Planen & Reflektieren" durchführen zu können.

Die Visualisierung des Wissensnetzes entspricht keinem eigenen Prozessschritt, sondern ist eine Funktion, die in verschiedenen Prozessschritten benötigt wird. Die typisierten Tags machen die Realisierung verschiedener Visualisierungen möglich bzw. erweitern Visualisierungen um neue Funktionen. Die Visualisierung des Wissensnetzes kann die Lernenden beim Wiederfinden, Reflektieren und Erkennen von Zusammenhängen unterstützen. Dabei gibt es allerdings nicht eine Darstellungsform, die für alle Anwendungsfälle gleichermaßen gut ist. Verschiedene Darstellungsformen haben gewisse Stärken und Schwächen. Des Weiteren ist für alle Darstellungsformen wichtig, dass eine Unterscheidung von Inhalten des Lernenden und anderen Lernenden erfolgt und fremde Inhalte bei Bedarf ausgeblendet werden können. Im Folgenden werden verschiedene Darstellungsformen erläutert, die im Szenario des Ressourcen-basierenden Lernens sinnvoll eingesetzt werden können:

- **Graph:** Das Wissensnetz besteht aus Knoten (Ressourcen, Tags) und Kanten (Relationen), die sich als Graph darstellen lassen.
 - **Überblick und Details:** Dieser Graph soll sowohl einen Überblick auf das Wissensnetz bieten als auch Details darstellen, wenn in das Wissensnetz hinein gezoomt wird. Die Inhalte eines selektierten Knotens (z.B. Snippets) sollen bei Bedarf eingeblendet werden können.
 - **Browsen:** Neben dem Zoomen soll auch eine Möglichkeit zum Browsen von Knoten zu Knoten angeboten werden, mit der sich benachbarte Knoten ein- und ausblenden lassen.
 - **Filtern:** Es sollen verschiedene Filter gesetzt werden können, z.B. Filter, die bestimmte Tag-Typen ein- und ausblenden. Die Filter können auch auf den Metadaten der Knoten ansetzen, z.B. Ausblenden aller Ressourcen, die vor einem bestimmten Datum gespeichert wurden.
- **Tabelle:** Die Knoten des Wissensnetzes lassen sich in Tabellen auflisten. Spalten können hierbei die Eigenschaften der Knoten sein (z.B. Bezeichner, Erstellungsdatum und Beschreibung).
- **Kalender:** Werden für Ziele Deadlines oder Bearbeitungszeiträume angegeben, lassen sich diese in einer Kalenderansicht darstellen. Ebenso kann der Zeitpunkt der Erstellung von Ressourcen und Tags in einem Kalender dargestellt werden und so als eine zeitliche Erinnerungshilfe dienen, z.B. wenn gleichzeitig im Kalender auch Termine eingetragen sind.
- **Karte:** In externen Quellen können Geo-Daten für Orts-Tags abgerufen werden, so lassen sich Ressourcen, die mit einem Ort getaggt wurden, auf einer geographischen Karte anzeigen.
- **Spickzettel:** Werden Kommentare und Zusammenfassungen von Ressourcen im Wissensnetz gepflegt, kann die Darstellung dieser Kommentare als geordnete Liste (ähnlich eines Spickzettels) beim Lernen dieser Inhalte helfen.
- **Schreibtisch:** Bei der Nutzung der Wissensnetzinhalt kann es hilfreich sein, die Ressourcen wie auf dem eigenen Schreibtisch räumlich anzuordnen und Stapel zu bilden. Dies ist im Umfeld der Wissensvisualisierung u.a. als Desktop-Metapher bekannt. Im Gegensatz zum physikalischen Schreibtisch lässt sich der Zustand des virtuellen Schreibtisches abspeichern und bei Bedarf als Erinnerungshilfe neu laden.

7.2 Technische Anforderungen und Design

Die funktionalen Anforderungen und das Konzept stellen gewisse Anforderungen an die technische Umsetzung. Diese betreffen zum einen das Datenmodell und zum anderen die Anbindung an einen Webbrowser. Das Datenmodell muss die sich aus dem semantischen Tagging ergebenden Anforderungen erlauben, insbesondere eine Erweiterbarkeit der Metadatenattribute durch die Nutzer selbst. Aus den Anforderungen und dem Konzept zum Speichern von Webseiten ergibt sich die Anforderung der Anbindung an einen Webbrowser.

7.2.1 Datenmodell

In diesem Abschnitt wird das Basisschema des semantischen Netzes beschrieben. Des Weiteren wird erläutert, wie Erweiterungen des Schemas durch die Lernenden realisiert werden können. Bei der Modellierung für das Wissensnetz müssen bestimmte Designentscheidungen getroffen werden, auf die im entsprechenden Abschnitt näher eingegangen wird.

Basisschema

Im semantischen Netz sollen Ressourcen, Tags und Relationen gespeichert werden. In Abbildung 7.1 ist das Basisschema dargestellt. Jede Ressource und jedes Tag entspricht einem Knoten im Netz, in der Abbildung als "Ding" bezeichnet, und wird über eine systemweit eindeutige ID identifiziert. Des Weiteren hat jedes Ding, wie in der funktionalen Analyse gefordert, einen Bezeichner, einen Erstellungszeitpunkt und einen Bearbeitungszeitpunkt. Ressource erbt diese Eigenschaften. Zusätzlich können für jede Ressource eine URL, eine Beschreibung, ein Screenshot und mehrere Kommentare gespeichert werden. Ein untypisiertes Tag hat keine Eigenschaften, außer die, die er von Ding erbt. Für die typisierten Tags im Basisschema kann eine Beschreibung gespeichert werden. Ziele und Themen können die in der Abbildung aufgelisteten Eigenschaften haben. Ressourcen können über Relationen mit Tags verknüpft werden. Tags können auch untereinander verknüpft werden. Ressourcen und Tags werden vom Nutzer erstellt. Außerdem kann jedes Ding durch den Nutzer erweitert werden. Für einen Nutzer kann ein Nickname, Vorname und Familienname, der Registrierungszeitpunkt, der letzte Login und seine Lerngruppen gespeichert werden.

Erweiterbarkeit des Datenmodells

Das Basisschema kann durch den Nutzer erweitert werden. In Abbildung 7.2 sind zwei Erweiterungen dargestellt, die die Ressource erweitern. Bei Bedarf kann der Nutzer für Ressourcen mit der "Bibtex"-Erweiterung den Titel, den Buchtitel und das Jahr und mit der "Bewertung"-Erweiterung einen Kommentar für Qualität und Relevanz ergänzen. Auch gleichzeitig beide Erweiterungen für eine Ressource zu nutzen, ist möglich. Die Erweiterungen sind ebenso für Tags möglich.

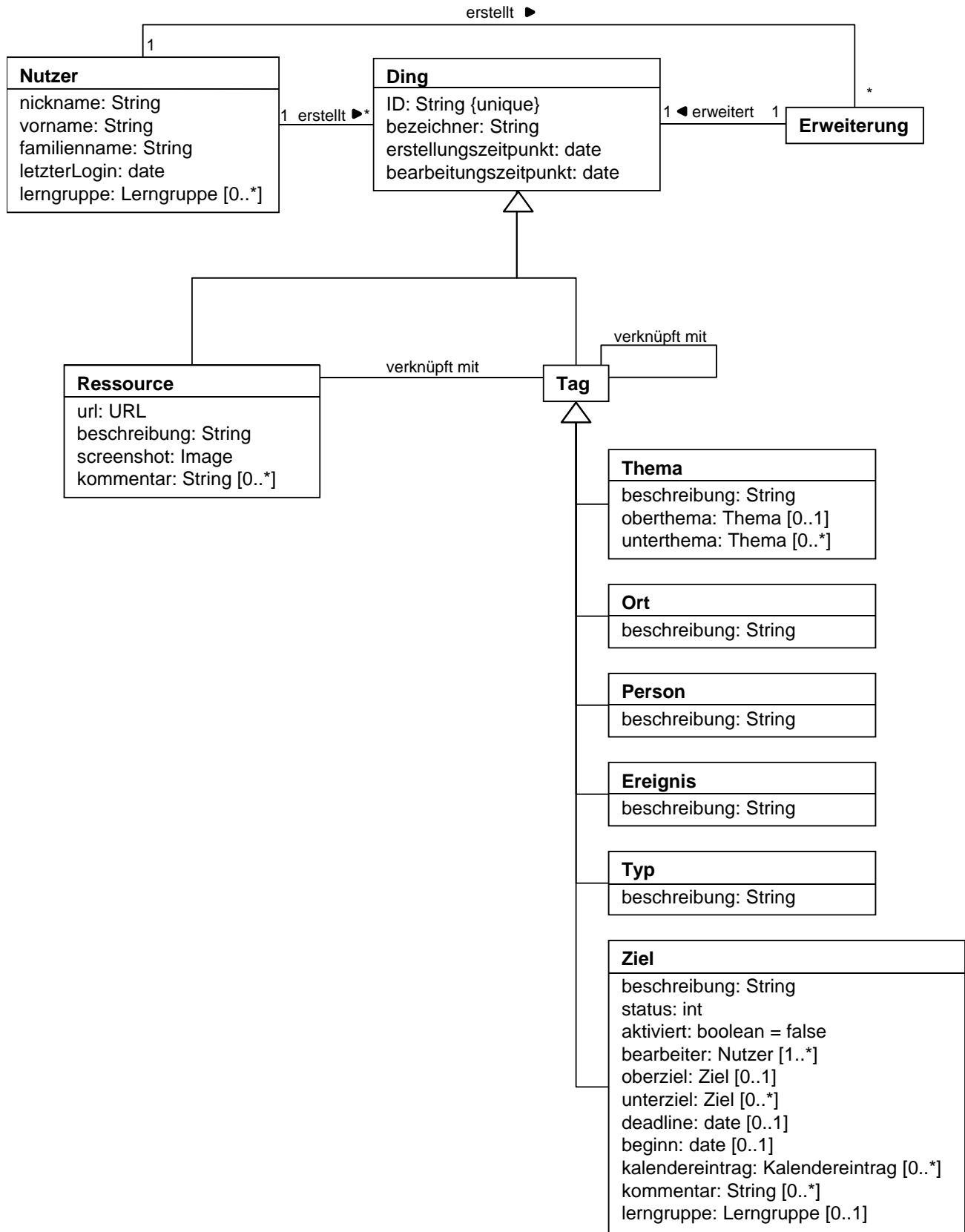


Abbildung 7.1: Basisschema

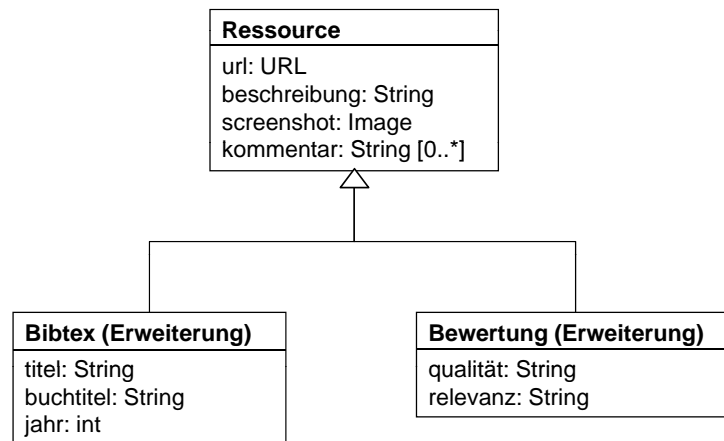


Abbildung 7.2: Erweiterungen im Basisschema

Designentscheidungen für die Modellierung

Bei der Modellierung des semantischen Netzes müssen teilweise spezielle Entscheidungen getroffen werden. Im Folgenden sind exemplarisch zwei Designentscheidungen graphisch dargestellt. In den Abbildungen sind die Knoten als Rechteck und die Eigenschaften als Ellipse dargestellt.

Mehrfaches Speichern einer Webseite

Eine Entscheidung betrifft das mehrfache Speichern einer Webseite. Da für eine Webseite mehrere unterschiedliche Snippets wichtig sein können, sollte eine Webseite mehrfach im Wissensnetz als Ressource gespeichert werden können. Da die URL einer Webseite nicht die eindeutige ID im Wissensnetz ist, sondern eine systemweit eindeutige ID erstellt wird, ist das Speichern der gleichen URL problemlos möglich.

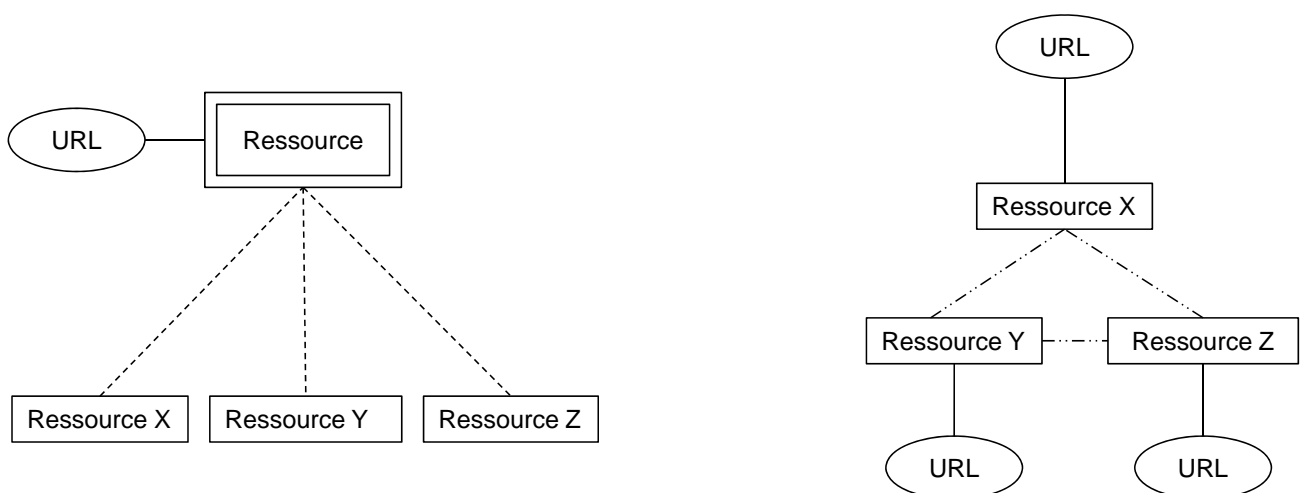


Abbildung 7.3: Mehrfaches Speichern einer Webseite: Ersatz-Ressourcenknoten (links) vs. verbundene Ressourcenknoten (rechts)

Für die Realisierung von Empfehlungen und der Pflege von Wissensnetzen (vgl. Konzept im vorhergehenden Abschnitt) ist es aber sinnvoll, die Information, dass es sich um Ressourcen einer identischen Webseite handelt, zu speichern. Dazu gibt es zwei Varianten der Modellierung.

Links in der Abbildung 7.3 ist eine Modellierungsvariante dargestellt, die einen zusätzlichen Ressourcenknoten (doppeltes Rechteck) einführt, um die URL und weitere gemeinsame Eigenschaften in diesem Knoten zu speichern. Die Eigenschaften, die sich für die Ressourcen unterscheiden, wie die Snippets, werden in jeweils einem Ressourcenknoten (Ressource X, Ressource Y und Ressource Z) gespeichert. Die einzelnen Ressourcenknoten erben die Eigenschaften des zusätzlichen Ressourcenknotens. Diese Variante ist dann sinnvoll, wenn sich die Ressourcen kaum voneinander unterscheiden und die identischen Eigenschaften nicht mehrfach gespeichert werden sollen.

Unterscheiden sich die Eigenschaften bis auf die URL ist die Variante rechts in der Abbildung besser. Für jede Ressource wird ein Knoten erstellt, in dem alle Eigenschaften gespeichert werden, auch die URL. Die Knoten selbst können über spezielle Relationen verbunden werden (gestrichelt-gepunktete Linien), um deren Zusammenhang explizit zu speichern.

Globales Wissensnetz

Das semantische Wissensnetz ist ein systemweit gültiges Wissensnetz. Daher muss, wie im Basisschema durch die Relationen "Nutzer erstellt Ding" bereits angedeutet, für jeden Knoten eine solche Relation erstellt werden, um die Knoten zu den Nutzern zuordnen zu können. Denn jeder Nutzer soll trotz der systemweiten Gültigkeit sein persönliches Wissensnetz aufbauen können. Eine weitere Designentscheidung ist hinsichtlich der Notwendigkeit zu treffen, den Benutzern das Erstellen unterschiedlicher Attribute zu identischen Ressourcen zu erlauben. Zwei Modellierungsvarianten sind in Abbildung 7.4 dargestellt. In

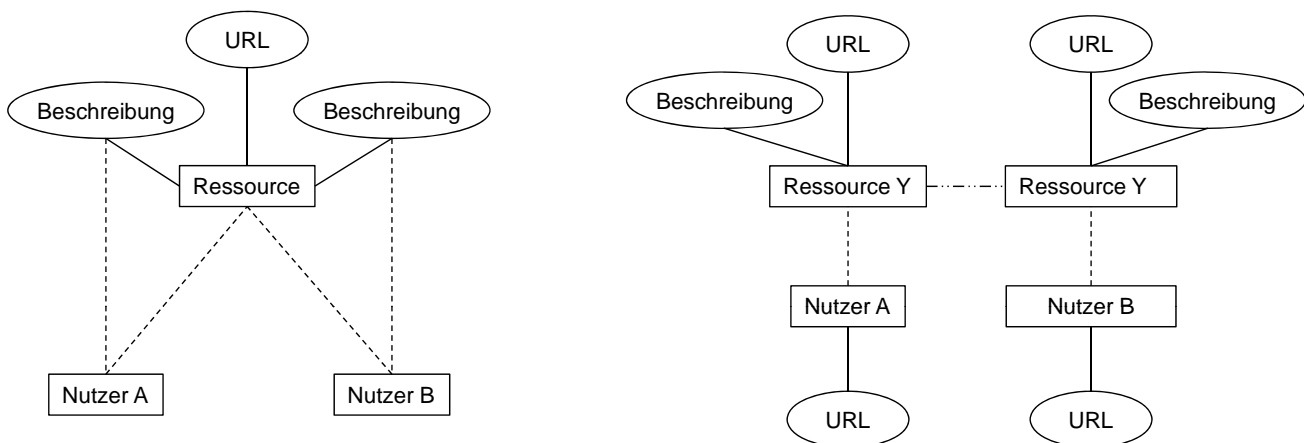


Abbildung 7.4: Global genutztes Wissensnetz: Gemeinsam genutzte Knoten (links) vs. verbundene Knoten (rechts)

der linken Variante wird eine Ressource von zwei Nutzern gespeichert. Jeder Nutzer will eine eigene Beschreibung ergänzen. Deshalb muss es in der Variante eine "erstellt"-Relation (gestrichelte Linie) vom Nutzer zum Knoten und vom Nutzer zur Eigenschaft (in diesem Falle zur Beschreibung) geben. In der Variante rechts wird für jeden Nutzer ein eigener Knoten angelegt, der mit dem Nutzer verknüpft ist, und auf Grund der Ähnlichkeit wird eine Relation zwischen den beiden Knoten (gestrichelt-gepunktete

Linie) erstellt. Die Ähnlichkeit kann einerseits vom Nutzer selbst festgestellt werden oder auch automatisch vom System durch die in der Anforderungsanalyse genannten Mapping-Verfahren (siehe Abschnitt 7.1).

7.2.2 Anbindung an den Webbrowser

Wie in der funktionalen Anforderungsanalyse beschrieben, ist der Webbrowser ein sehr wichtiges Werkzeug, um neue Ressourcen zu sammeln. Die Anbindung an den Webbrowser ist daher gefordert, um den manuellen Aufwand beim Speichern der Ressourcen gering zu halten.

Für die Realisierung der Anbindung gibt es unterschiedliche Lösungen: eine eigenständige Applikation, eine Webanwendung und ein Plug-in im Webbrowser. Eine eigenständige Applikation hat den Vorteil, dass sie sehr frei umgesetzt werden kann im Sinne der Mächtigkeit und im Sinne der Abhängigkeiten zu anderen Applikationen. Sie hat allerdings den Nachteil, dass eine Anbindung an den Webbrowser schwierig und durch Sicherheitsbeschränkungen nur eingeschränkt möglich ist.

Ein Plug-in, integriert im Webbrowser, hat den Vorteil, dass ein erweiterter Zugriff auf den Webbrowser besteht und die Nutzer nicht zwischen verschiedenen Anwendungen hin und her wechseln müssen, um Ressourcen im Wissensnetz zu speichern oder gespeicherte Ressourcen im Webbrowser zu laden. Ein Plug-in wiederum hat den Nachteil, dass der Platz zur Darstellung der Inhalte des Wissensnetzes begrenzt ist.

Eine Webanwendung hat für die Darstellung die Größe des Browser-Fensters zur Verfügung, aber einen ähnlich eingeschränkten Zugriff auf den Webbrowser wie eine eigenständige Applikation.

Daher ist die bessere Lösung eine Kombination aus Plug-in und Webanwendung, denn reicht die Darstellungsgröße nicht aus, kann vom Plug-in zur Webanwendung gewechselt werden, ohne den Webbrowser verlassen zu müssen.

7.3 Zusammenfassung

Ausgehend vom Modell des Ressourcen-basierten Lernens wurden Funktionen erarbeitet, die der Unterstützung der Lernenden beim Ressourcen-basierten Lernen dienen. Die Funktionen decken alle identifizierten Prozessschritte des Modells ab. Bei der Erarbeitung des Konzeptes wurden vor allem Aspekte beachtet, die für das persönliche Ressourcenmanagement wichtig sind. Funktionen, die eng mit der Community, also verschiedenen Lernenden, zusammenhängen, sind nur teilweise und knapp beschrieben. Sie sind nicht Gegenstand dieser Arbeit, werden aber in dem BMBF-Verbundprojekt "CROKODIL" weiter aus- und fortgeführt.

Der Aufbau eines Wissensnetzes durch die Lernenden erfüllt die in [92] genannten Anforderungen an ein effizientes Lernen wie folgt:

- Lernen ist ein aktiver Prozess. → Der Lernende muss beim Ressourcen-basierten Lernen selbst aktiv sein.
- Lernen ist ein selbstgesteuerter Prozess. → Das Konzept bietet Funktionen, die die Prozesse des selbstgesteuerten Lernens anregen können.
- Lernen ist ein konstruktiver Prozess. → Wissen wird auf Grund des typisierten Taggings in vorhandene Wissensstrukturen eingebaut.

-
- Lernen ist ein situativer Prozess. → Der Erwerb von Wissen ist an einen spezifischen Kontext gebunden. Die Ziele bilden den Kontext für die Recherche und das Lernen.
 - Sozialer Prozess → Die Konstruktion von Wissen geschieht häufig durch Kommunikation und Zusammenarbeit. Das Konzept sieht Funktionen zur Kommunikation (im Sinne der Weitergabe von Wissen) und zur Zusammenarbeit (d.h. die Bildung von Lerngruppen) vor.

Die Voraussetzungen für die Qualitätsverbesserung des Lernprozesses sind durch das Konzept gegeben. Im Kapitel 9 wird näher auf die tatsächlich erreichte Qualitätsverbesserung des Lernprozesses eingegangen. Eine Untersuchung der Qualitätsverbesserung der Lernergebnisse liegt außerhalb dieser Arbeit.

8 Implementierung und Proof-of-Concept

Dieses Kapitel beschreibt die Implementierung des Konzeptes aus Kapitel 7. Dem implementierten Werkzeug wurde der Name "ELWMS.KOM" gegeben, der kurz für "E-Learning-Wissens-Management-System" steht und "KOM" ist die Abkürzung für das Fachgebiet "MultimediaKOMmunikation", an dem das Werkzeug entwickelt wurde. Im ersten Teil des Kapitels wird die Architektur des Werkzeugs beschrieben. Die implementierten Funktionen, die in der Anforderungsanalyse festgestellt und im Konzept beschrieben wurden, werden im zweiten Teil dieses Kapitels erläutert.

8.1 Komponenten und Architektur

ELWMS.KOM besteht aus drei verschiedenen Software-Komponenten: einem Client, einem Web-Service und einem Backend. Diese Komponenten werden nachfolgend einzeln beschrieben. Anschließend wird die Gesamtarchitektur von ELWMS.KOM erläutert.

8.1.1 Backend: K-Infinity

K-Infinity ist eine Technologieplattform der Firma intelligent views¹ zur Verwaltung von Wissensnetzen. Sie besteht aus verschiedenen Komponenten, u.a. aus:

- einer proprietären Datenbank, die das Wissensnetz speichert,
- Schnittstellen, die den Zugriff auf das Wissensnetz realisieren,
- Werkzeugen zur Modellierung des Wissensnetzes und
- einem Web-Portal als Frontend für die Visualisierung und die Bearbeitung von Individuen des Wissensnetzes durch die Nutzer.

Durch die Implementierung der Komponenten in Java und Smalltalk ist K-Infinity plattformunabhängig. Ein Werkzeug zur Modellierung des Wissensnetzes ist der Knowledge-Builder. Dieser besteht aus dem K-Organizer und dem Graph-Editor. Mit beiden Werkzeugen können Wissensnetzknotten und Relationen erzeugt, bearbeitet und gelöscht werden, wobei der K-Organizer das Wissensnetz baumartig und der Graph-Editor das Wissensnetz als Graph darstellt. Wissensnetzknotten werden in K-Infinity in Begriffe, Individuen und Erweiterungen unterschieden. Für jeden Wissensnetzknotten können Attribute und Relationen definiert werden. Diese modellierten Knoten, Attribute und Relationen bilden das Schema des Wissensnetzes. Dieses Schema wird in der Regel von Wissensmodellierungs-Experten modelliert. Die Individuen bilden dann den tatsächlichen Inhalt des Wissensnetzes. Sie können zum einen durch die Experten oder die Nutzer des Systems erstellt werden oder zum anderen durch automatische Importmechanismen.

Abbildung 8.1 zeigt die Modellierung des Basisschemas aus dem Konzept für das Werkzeug ELWMS.KOM im Graph-Editor. Begriffe werden durch runde Kreise und Attribute durch Dreiecke dargestellt. Die Kreise mit inneren gefüllten Kreisen zeigen Begriffe, zu denen Individuen erstellt werden

¹ <http://www.i-views.de>, Zuletzt abgerufen am 20.02.2011

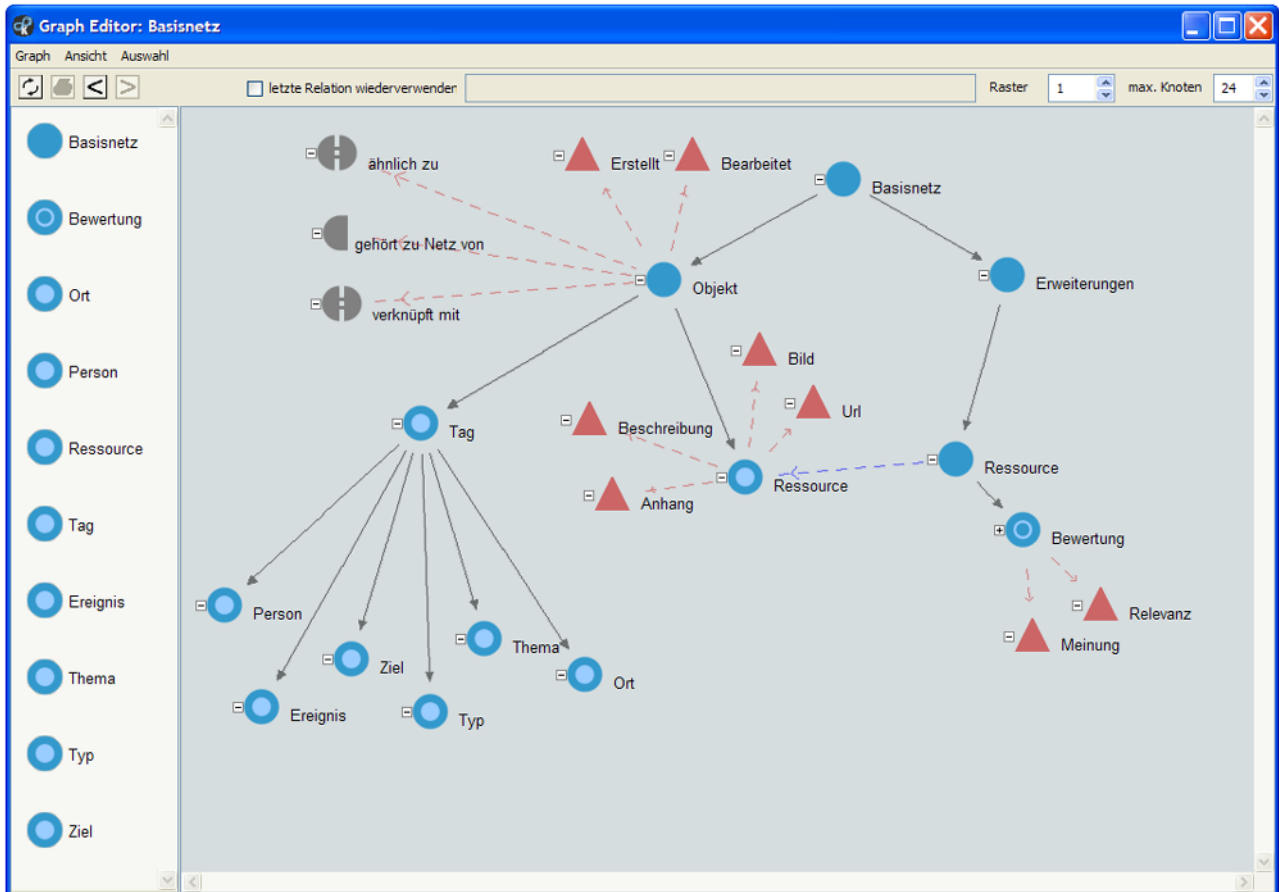


Abbildung 8.1: Modelliertes Wissensnetz in K-Infinity

dürfen. Die durchgezogenen Linien mit Pfeilspitze entsprechen Vererbungsrelationen. Das Wissensnetz benötigt einen Wurzelknoten, in diesem Falle den Begriffsknoten "Basisnetz". Jedes Element (Begriffe, Individuum,...) in K-Infinity muss die Attribute "Name" und "ID" besitzen, wobei die im Wissensnetz eindeutige ID automatisch vom System bei Erstellung des Elements vergeben wird. Der Begriff "Objekt" entspricht dem "Ding" im Basisschema des Konzeptes. Es definiert die Attribute "Erstellt" und "Bearbeitet" und vererbt diese an alle Unterbegriffe, in diesem Fall an "Tag" und "Ressource". Unterbegriffe von "Tag" sind die typisierten Tags, deren Attribute in der Abbildung ausgeblendet sind. Neben den geerbten Attributen sind für Ressourcen deren definierte Attribute "Url", "Beschreibung", "Bild" und "Anhang" eingeblendet. Die gestrichelte Linie von diesem Kreis führt in der Abbildung nach rechts zu dem Kreis, der ebenfalls den Namen "Ressource" trägt, einem Unterbegriff von dem Begriff "Erweiterungen". Alle Unterbegriffe dieses Begriffs stellen Erweiterungen dar, sie werden als Kreise mit einer inneren Kreislinie gezeichnet. In der Abbildung ist nur die Erweiterung "Bewertung" für Ressourcen mit den Attributen "Meinung" und "Relevanz" eingeblendet.

Die Relationen werden im Graph-Editor durch Halbkreise (asymmetrische Relationen) und doppelte Halbkreise (symmetrische Relationen) dargestellt. Die asymmetrische Relation "gehört zu Netz von" verknüpft Objekte des Wissensnetzes mit Benutzern, dessen Begriffsknoten aus Gründen der Übersichtlichkeit ausgeblendet wurde. Die anderen beiden Relationen "ähnlich zu" und "verknüpft mit" haben Objekte sowohl als Relationsquelle als auch als Relationsziel.

Die Nutzer des Systems erstellen die Individuen jedoch nicht im Knowledge-Builder, sondern im Knowledge-Portal. Dies ist eine Web-Applikation, die automatisch Webseiten für die Darstellung und Eingabemasken zur Bearbeitung der Knoten des Wissensnetzes generiert. Im Knowledge-Portal können neue Individuen erstellt werden, das Schema aber kann nicht verändert werden.

8.1.2 Web-Service

K-Infinity bietet mit der KEM-API eine allgemeine Schnittstelle für den Zugriff zum Wissensnetz. Mit der KEM-API kann man daher nur Daten über Knoten und Kanten abrufen. Deshalb wird für ELWMS.KOM ein Web-Service benötigt, der die abstrakten Daten entsprechend des Konzeptes von Ressourcen und Tags für den Zugriff des Clients aufbereitet. Der Web-Service erfüllt damit eine Art Mittlerfunktion zwischen Client und Backend. Im Abschnitt 8.1.4 wird die Funktion des Web-Service im Zusammenspiel mit den beiden anderen Komponenten erläutert.

8.1.3 Add-on

In der Anforderungsanalyse wurde gefordert, dass das Frontend von ELWMS.KOM direkt in den Webbrowser integriert wird, da der Webbrowser ein wichtiges Werkzeug beim Ressourcen-basierten Lernen ist. Daher gibt es neben dem Knowledge-Portal von K-Infinity ein weiteres Front-End - das Add-on. Es ist als Java-Applet implementiert, das in den Mozilla Firefox-Browser integriert ist, der eine entsprechende API zur Integration von Add-ons anbietet. Die Hauptfunktionalität dieses Add-ons ist die Unterstützung der Recherche im Webbrowser.

8.1.4 Gesamtarchitektur

Die Gesamtarchitektur ist als Client-Server-Architektur realisiert. Sie ist in Abbildung 8.2 schematisch dargestellt. Die gestrichelte vertikale Linie teilt die Architektur in eine Client-seitige Komponente und die Server-seitigen Komponenten.

Werden vom Client Daten aus dem Wissensnetz benötigt, wird eine Nachricht über SOAP an den Web-Service gesendet. SOAP stand ursprünglich für Simple Object Access Protocol, diese Bezeichnung wird allerdings seit Version 1.2 [52] der W3C-Recommendation nicht mehr verwendet. Der Web-Service übersetzt diese Nachricht in die Sprache der KEM-API und bereitet deren zurückgegebenen Daten auf und schickt diese an den Client zurück. Ebenso erfolgt der umgekehrte Weg, wenn der Client Daten im Wissensnetz speichern soll. Der Client schickt die Daten an den Web-Service, dieser wandelt die empfangenen Daten wiederum in Daten um, die von der KEM-API verstanden werden. Im nächsten Abschnitt ist der Ablauf des Nachrichtenaustausches anhand einer Abbildung dargestellt.

Des Weiteren gibt es eine direkte Verbindung vom Client zum Knowledge-Portal, das der Nutzer bei Bedarf mittels Client-Funktion im Webbrowser öffnen kann.

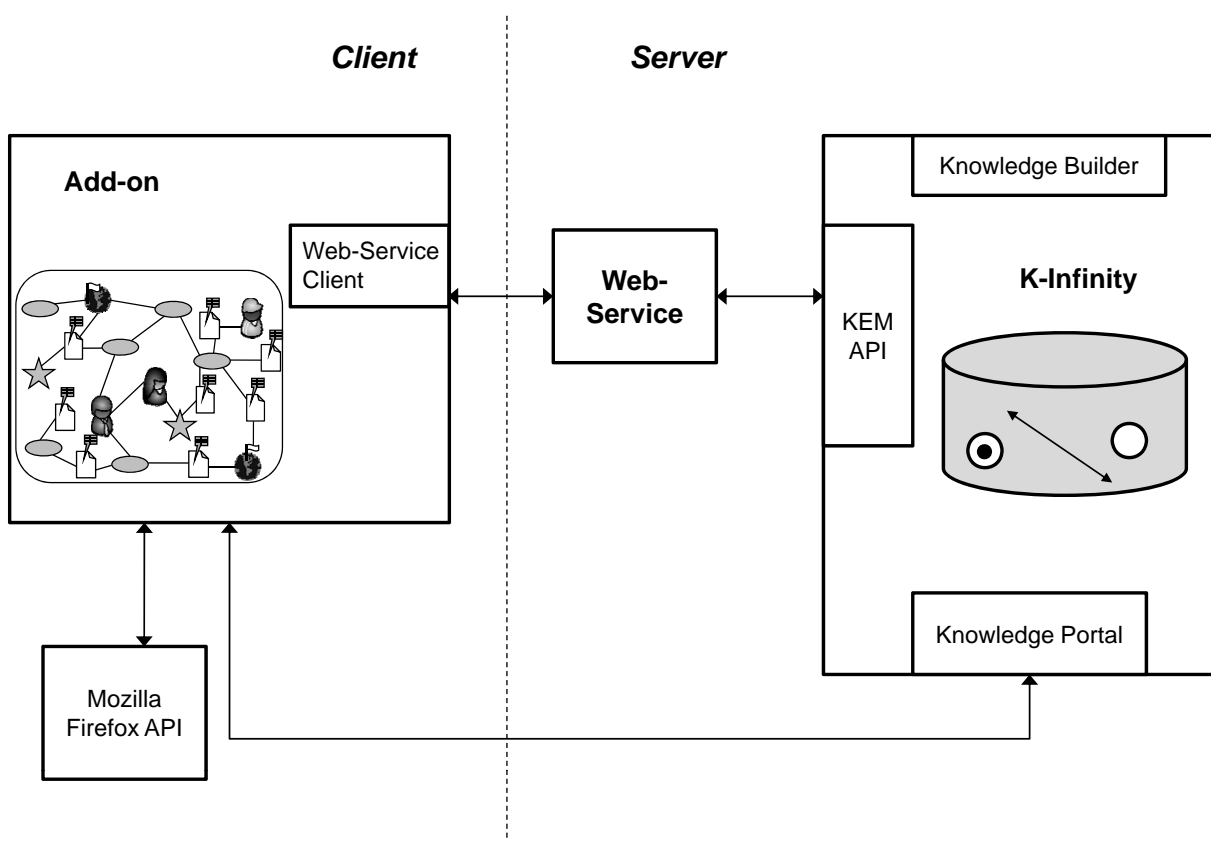


Abbildung 8.2: Gesamtarchitektur von ELWMS.KOM

Beispielanfrage

In Abbildung 8.3 ist der Nachrichtenaustausch zwischen den Komponenten anhand einer beispielhaften Abfrage sequentiell dargestellt. Die drei Objekte entsprechen den jeweiligen Objekten in den drei Komponenten.

Der Client fordert mittels SOAP-Nachricht, die den Methodenaufruf "holeKnoten" mit den Parametern "Ressource" und dem Nutzerobjekt enthält, beim Web-Service die Ressourcen des Nutzers an und erwartet eine Liste der entsprechenden Knoten aus dem Wissensnetz zurück. Der Web-Service ruft nach Erhalt dieser Nachricht die Daten aus dem Wissensnetz mittels KEM-API ab. Zuerst fordert er mit dem Methodenaufruf eine Liste der IDs aller Ressourcen des Nutzers an. Anschließend fordert er für jeden Knoten die Attribute und die Relationen an, indem er jeweils als Parameter die ID des Knotens übergibt. Dies macht er für jede ID in der erhaltenen Liste aller Knoten. Aus diesen Daten baut der Web-Service eine Liste aus Ressourcenknoten zusammen, die Attribute und Relationen enthalten, und schickt diese Liste an den aufrufenden Client mittels SOAP-Nachricht zurück. Auf die SOAP-Nachrichten soll an dieser Stelle nicht explizit eingegangen werden.

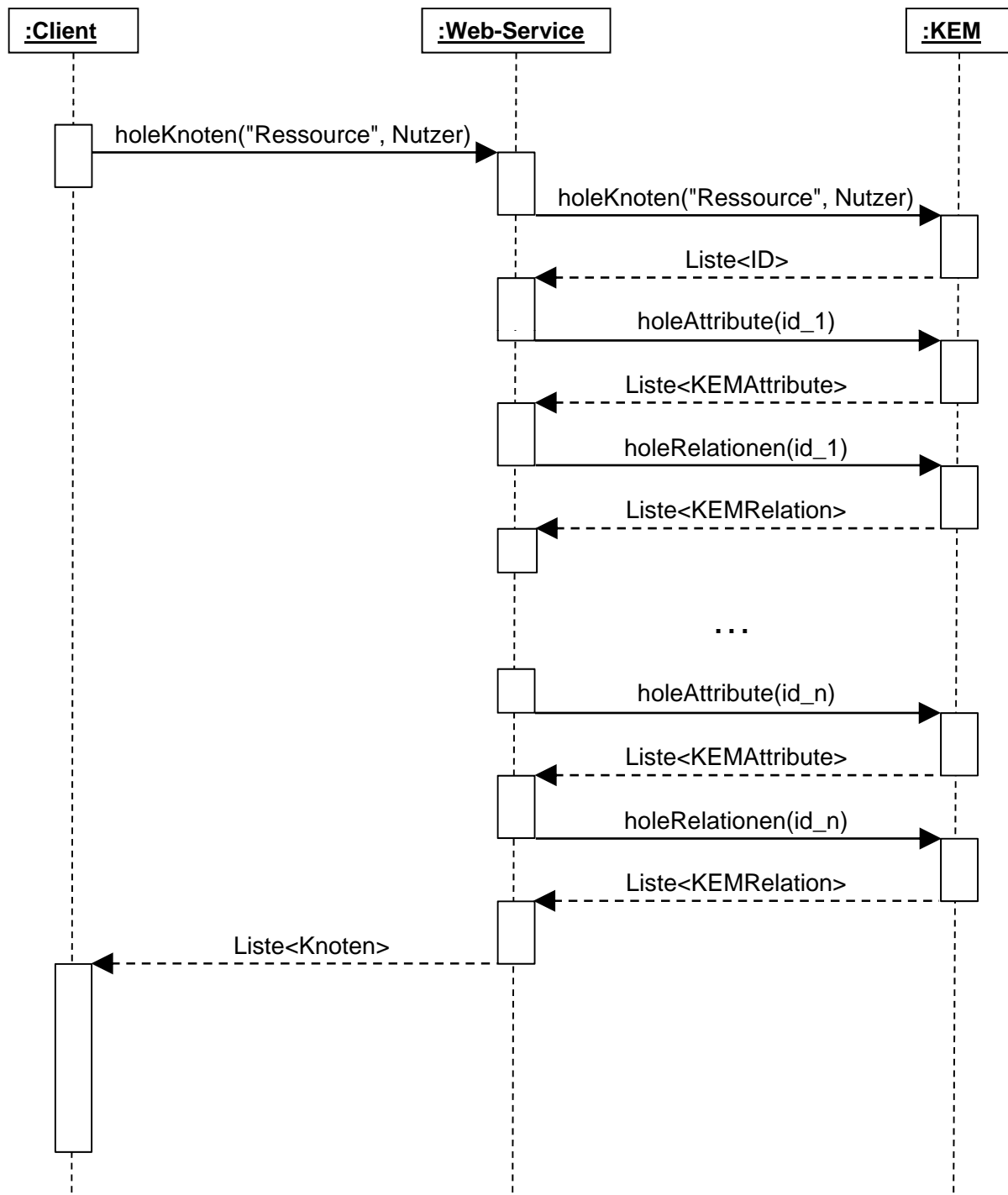


Abbildung 8.3: Beispiel von Datenaustausch zwischen Client und Server

Diese Aufteilung der Komponenten mit dem Web-Service als Vermittler-Komponente hat mehrere Vorteile:

- Die Details der KEM-API können vor dem Client verborgen werden, da dieser nur mit dem Web-Service kommuniziert, der auf den Client zugeschnittene Funktionen bereitstellt. Dadurch werden die Funktionsaufrufe des Clients vereinfacht.

-
- Die K-Infinity-Plattform wird von Zeit zu Zeit auf Grund von Weiterentwicklungen der Plattform aktualisiert. Durch das Verbergen der Funktionsaufrufe der KEM-API im Web-Service kann K-Infinity aktualisiert werden, ohne dass ein Austausch des Clients auf den Rechnern der Nutzer notwendig wird. Sogar der komplette Austausch von K-Infinity durch eine andere Software zur Verwaltung des Wissensnetzes ist möglich ohne den Client verändern zu müssen.
 - Der Datenaustausch via SOAP über das Netzwerk ist bedeutend langsamer als der Datenaustausch zwischen Web-Service und KEM-API, vor allem wenn deren Aufruf nicht über das Netzwerk geschehen muss, sondern auf dem gleichen Server stattfinden kann. Die Daten des Wissensnetzes müssen mit der KEM-API in vielen kleinen Funktionsaufrufen ausgetauscht werden. In Abbildung 8.3 sind nur ein Teil der Funktionsaufrufe abgebildet. Die Geschwindigkeit des direkten Datenaustauschs von KEM-API und Client wäre in dieser Form zu gering.

8.2 Funktionen

In diesem Abschnitt werden die Funktionen, die in ELWMS.KOM umgesetzt wurden, anhand von Screenshots erläutert. Die Funktionen wurden im Konzept in Kapitel 7 beschrieben und aus der Anforderungsanalyse abgeleitet. Die Beschreibung der Funktionen ist unterteilt in die Funktionen des Add-ons und des Portals.

8.2.1 Add-on

Das Add-on dient primär der Unterstützung der Recherche im Webbrowser und der damit unmittelbar verbundenen Funktionen des *Annotierens und Organisierens*, vgl. Konzept, und deren Einbindung in die Funktionen des *Planens und Reflektierens*. Darüber hinaus bietet es weitere Funktionen aus den Bereichen *Suchen, Nutzen und Weitergeben*.

Der Screenshot 8.4 zeigt das Add-on, das in den Webbrowser Mozilla Firefox integriert ist, genauer in die Sidebar des Webbrowsers. Bei Bedarf lässt sich die Sidebar einklappen. Die Funktionen des Add-ons sind auf zwei Tabulatoren verteilt, um die Übersichtlichkeit der Funktionen und Informationen in der Sidebar zu gewährleisten. Der erste Tabulator vereint die Funktionen, die der Lernende im Zusammenhang mit der Recherche im Web verwendet, der zweite Tabulator hingegen die Funktionen zur Suche im Wissensnetz.

Anzeige von Informationen zur aktuell geöffneten Webseite

Die Sidebar ist im ersten Tabulator in drei Bereiche unterteilt:

- **Informationen zur aktuellen Webseite:** In diesem Bereich werden, sofern die im Browser geöffnete Webseite bereits im Wissensnetz gespeichert ist, Informationen zu der Ressource aus dem Wissensnetz angezeigt. In diesem Falle erscheint die Nachricht, dass die aktuell geöffnete Webseite bereits vom Nutzer im Wissensnetz gespeichert wurde. Wurde sie nicht vom Benutzer selbst, sondern einem anderen Benutzer ins Wissensnetz aufgenommen, wird dies ebenfalls mit einer anderen farblichen Darstellung angezeigt. Des Weiteren werden dort ähnliche Ressourcen empfohlen, falls

von den in Kapitel 4.4 beschriebenen Recommender-Verfahren Empfehlungen ermittelt wurden. Der Nutzer kann die empfohlenen Ressourcen unmittelbar entweder im Original im Webbrowser öffnen oder er kann die Ressource mit ihren Attributen im Knowledge-Portal im Webbrowser öffnen. Die Ressource dient ihm im letzteren Fall als Eintrittspunkt in das Wissensnetz.

- **Zuletzt gespeicherte Ressourcen:** Die zuletzt gespeicherten Webseiten werden als minimierte Screenshots der Webseiten im unteren Bereich der Sidebar präsentiert. Ein Nutzer kann sich somit beobachten, ob er aktuell Webseiten gespeichert hat und wieviele. Dies hilft dabei, den Nutzer bei der Reflektion der Recherche zu unterstützen.
- **Ziele:** Im Bereich Ziele finden sich verschiedene Funktionen zur Unterstützung des Prozessschrittes *Planen und Reflektieren*. Diese werden nachfolgend detaillierter beschrieben.

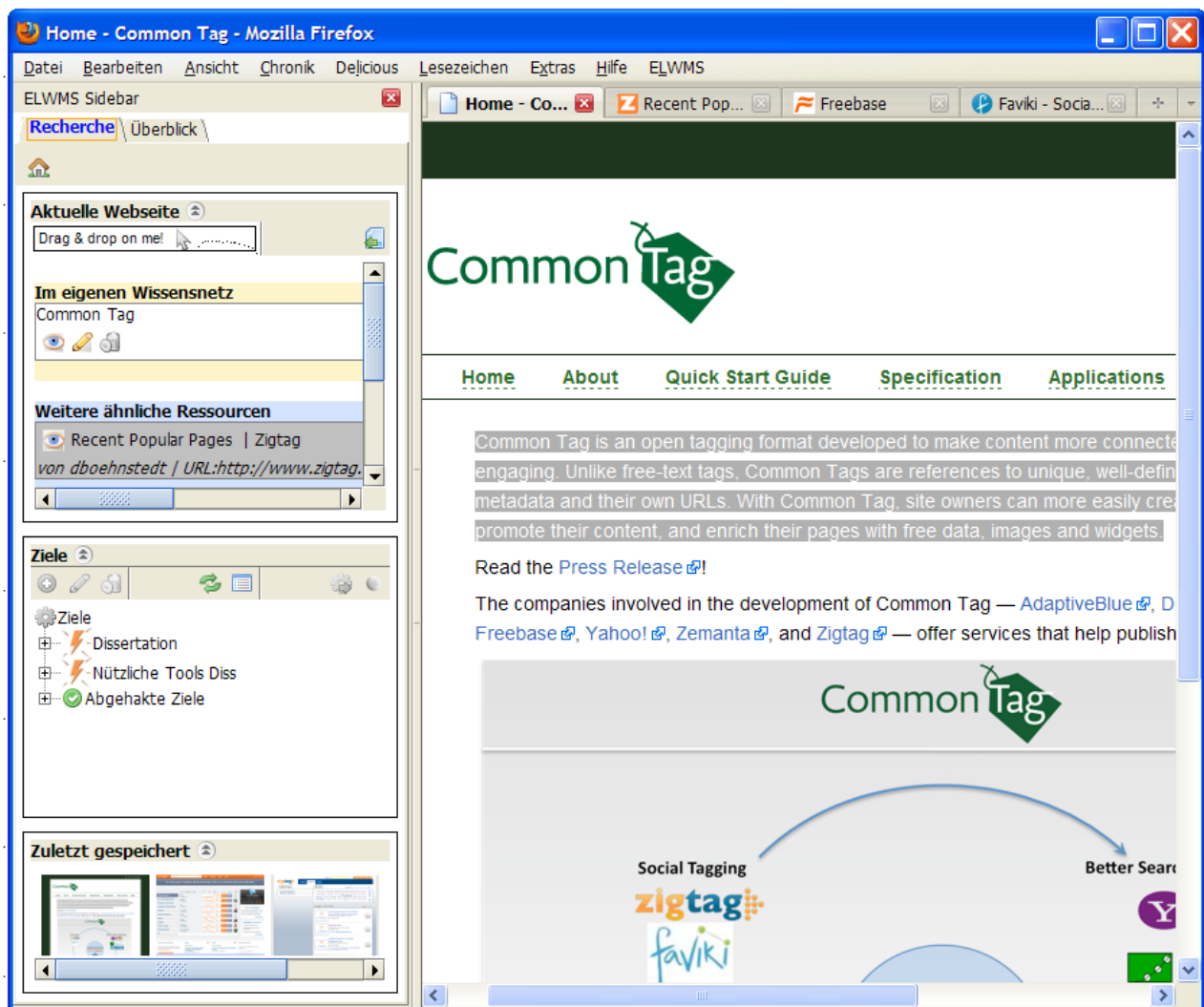


Abbildung 8.4: Sidebar "Recherche"

Speichern, Annotieren und Taggen von Ressourcen

Hat der Benutzer bei seiner Recherche im Web eine für ihn relevant erscheinende Webseite gefunden, kann er sie unmittelbar über das Add-on als Ressource ins Wissensnetz aufnehmen, d.h. speichern sowie annotieren und taggen. Der Benutzer löst den Vorgang des Speicherns einer Webseite als Ressource aus, indem er innerhalb der Ressource einen für ihn relevanten Abschnitt, das sogenannte Snippet, markiert und per Drag&Drop in ein Eingabefeld im oberen Bereich des Add-ons zieht. Darauf hin öffnet sich ein neues Eingabefenster in dem der Benutzer die Ressource annotieren und taggen kann. Die folgenden Attribute zur Beschreibung der Ressource werden dabei automatisch gefüllt, können vom Nutzer aber modifiziert werden:

- **Name:** In dieses Attribut wird der Titel der Webseite aus dem Header der Webseite aufgenommen.
- **Beschreibung:** Der vom Benutzer in der Webseite markierte Text, das Snippet, wird in das Attribut Beschreibung übernommen.
- **URL:** Die URL der Webseite wird in das entsprechende Attribut übernommen.
- **Erstellt:** Das aktuelle Datum und die Uhrzeit werden in das Attribut eingetragen.

The screenshot shows a window titled "Bearbeiten" (Edit) for a resource. It contains the following fields and sections:

- Resource:**
 - Name:** Resource-Based Learning: A Shift in the Roles of Teacher, Learner
 - Beschreibung:** The subtle shift in the teacher's role in resource-based learning is the single most significant factor in determining success in student learning through this approach, says this writer, who explains why.
 - Url:** <http://bul.sagepub.com/cgi/content/abstract/75/535/15>
 - Bearbeitet:** 23.02.2009 16:24
 - Erstellt:** 23.02.2009 16:24
- Neue Tags:**
 - ☒ Tags von allen Benutzern anzeigen
 - Tag:** [input field]
 - Thema:** [input field]
 - Person:** [input field]
 - Ort:** [input field]
 - Ereignis:** [input field]
 - Ziel:** [input field]
 - Sonstiges:** [input field]
- Erweiterung:**
 - Wissenschaftliche Ressource [dropdown menu]
 - Erweiterung hinzufügen [button]
- Relationen:**
 - verknüpft mit: [icon] Ressourcen-basiertes Lernen (Thema) [button: Entferne Verknüpfung]

At the bottom are buttons for "Speichern" (Save) and "Abbrechen" (Cancel).

Abbildung 8.5: Alter Tagging-Dialog

Abbildung 8.5 zeigt die erste Implementierung des Tagging-DIALOGs, der in der Studie "Seminar 1" zum Einsatz kam. Dieser Dialog enthielt für jeden Tag-Typ ein eigenes Eingabefeld, so dass der Nutzer vor der Eingabe des Tags in das richtige Eingabefeld springen musste. Aufgrund der Rückmeldung der Nutzer wurde dieser Tagging-Dialog angepasst und die Anzahl der Eingabefelder im neuen Tagging-Dialog auf

ein Eingabefeld reduziert. Abbildung 8.6 zeigt den neuen Tagging-Dialog bei Speichern einer Ressource. Im oberen Teil des Tagging-Dialogs befinden sich die Eingabefelder für die oben aufgeführten Attribute der Ressource.

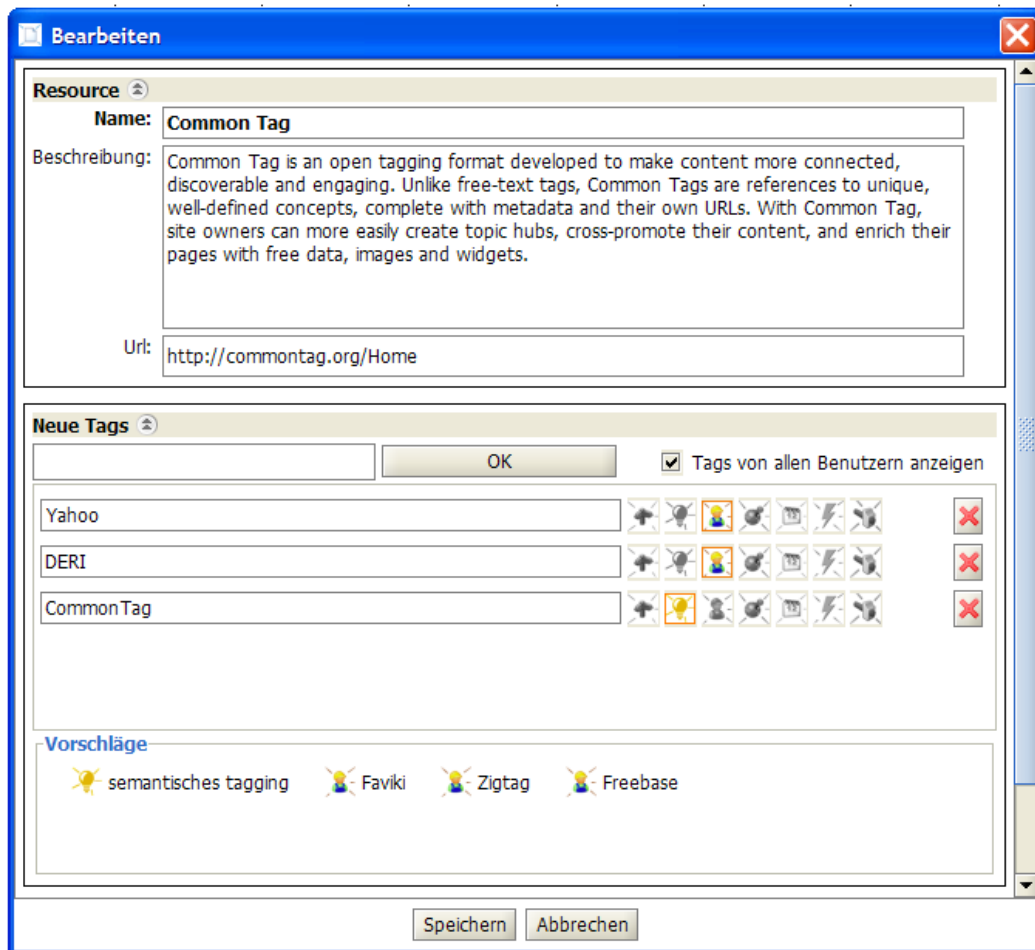


Abbildung 8.6: Tagging-Dialog

Im mittleren Bereich des Tagging-Dialogs befindet sich das Eingabefeld für neue Tags und Eingabefelder für bereits eingegebene Tags. Wird ein neues Tag eingetippt, schlägt die Autovervollständigungsfunktion Tags aus dem Wissensnetz des Nutzers und/oder anderer Nutzer vor. Anschließend kann das Tag typisiert werden. Die verschiedenen Tag-Typen werden durch Symbole dargestellt, die durch Anklicken als Tag-Typ übernommen werden. Im unteren Bereich des Tagging-Dialogs werden Tags inklusive Tag-Typ vorgeschlagen. Durch Klick können diese Tags übernommen werden, der Tag-Typ ist noch manuell veränderbar.

In ELWMS.KOM ist es, wie im Abschnitt Datenmodell in Kapitel 7.2.1 beschrieben, möglich, dass die Nutzer das Schema des Wissensnetzes erweitern und eigene Metadaten (Eigenschaften der Wissensnetz-knoten) definieren. Abbildung 8.7 zeigt im Hintergrund den Dialog zur Verwaltung der benutzerdefinierten Erweiterungen. Im Vordergrund ist der Dialog zur Erstellung/Bearbeitung einer Erweiterung abgebildet. Im dargestellten Fall wird die Erweiterung "Bewertung" erstellt, die zwei Attribute enthält. Für die nutzerdefinierten Eigenschaften kann angegeben werden, ob im Eingabedialog für die Eigenschaften beim Tagging mehrzeilige Textfelder oder einzeilige Textfelder verwendet werden sollen. Die

Verwendung der vom Benutzer definierten Erweiterungen der Metadaten im Rahmen des Tag-Dialogs zeigt die Abbildung 8.8. Im unteren Bereich können die Eingabefelder bei Bedarf eingeblendet und Informationen gespeichert werden.

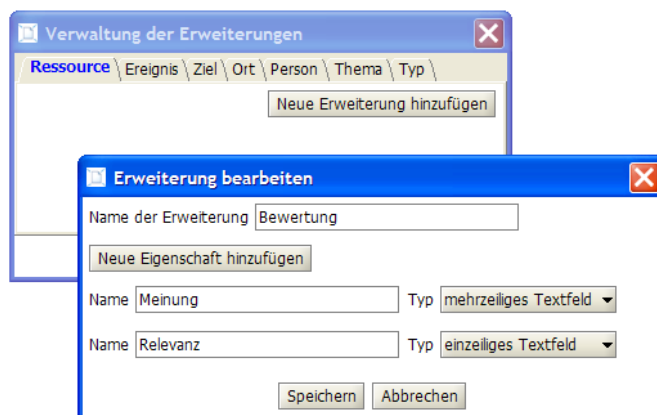


Abbildung 8.7: Erweiterungen definieren

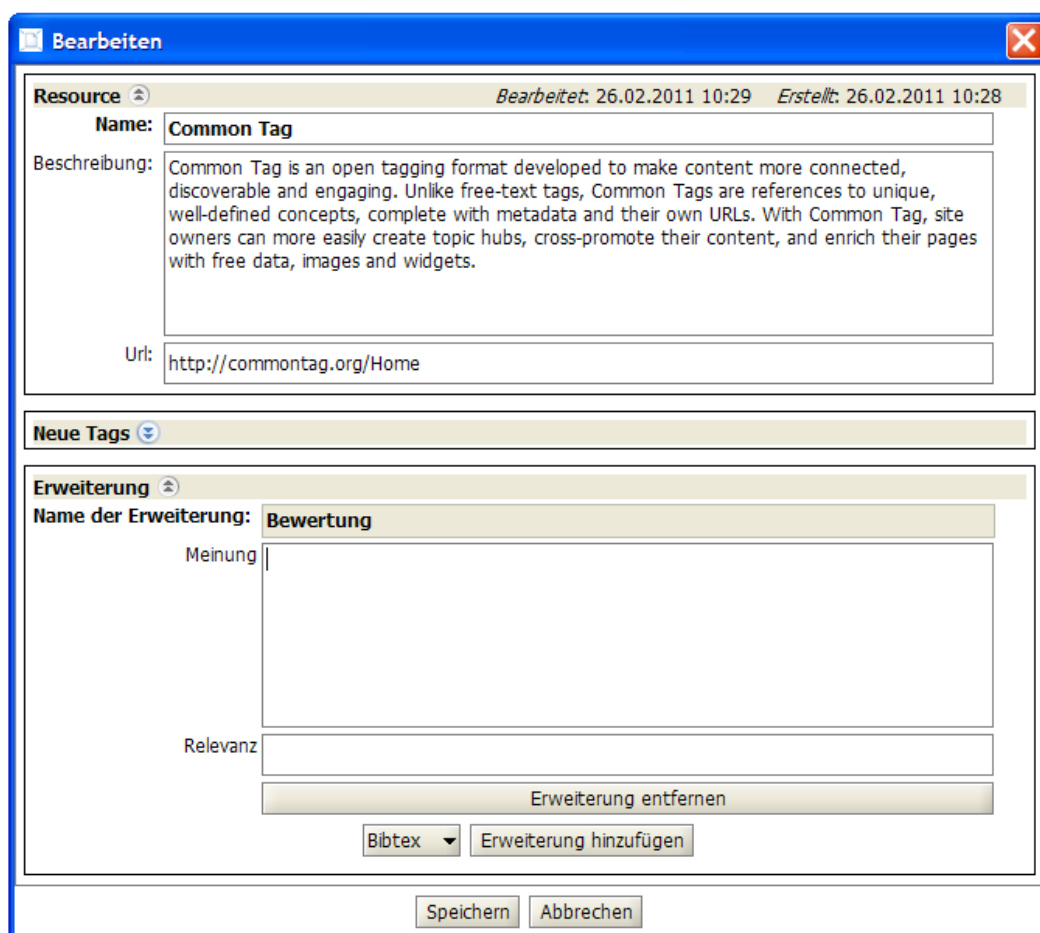


Abbildung 8.8: Erweiterung im Tagging-Dialog

Das Konzept für ELWMS.KOM sieht verschiedene Funktionen zur Unterstützung des Prozessschrittes *Planen und Reflektieren* vor. Zentrales Element ist die Zuordnung von Ressourcen zu Zielen. Die Zuordnung erfolgt grundsätzlich wie die Zuordnung von anderen Tags im Rahmen des Tagging-Dialogs.

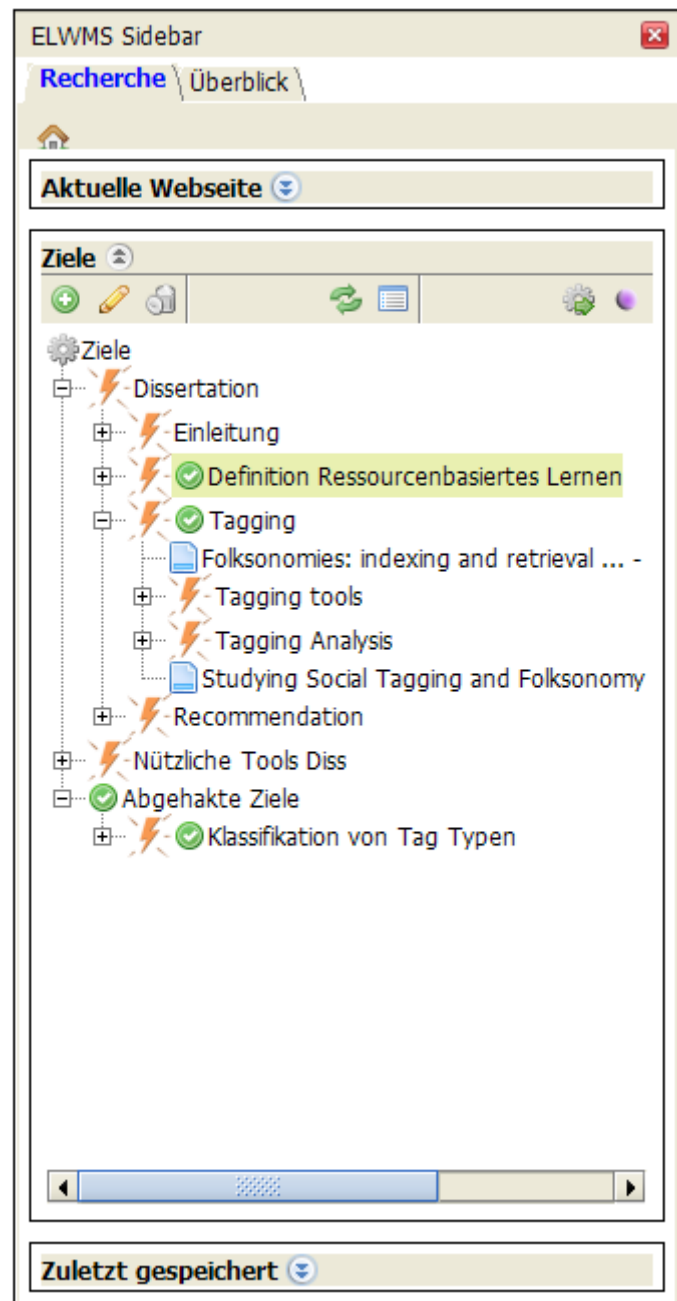


Abbildung 8.9: Zielmanagement

Da es sich bei den Zielen um ausgezeichnete Tags handelt, die von großer Bedeutung für das dem ELWMS.KOM zugrundeliegende didaktische Konzept sind und sie über eigene Attribute verfügen, exis-

tieren spezielle Funktionen, die im mittleren Bereich des Add-ons realisiert sind. Dieser Bereich der Ziele ist in Abbildung 8.9 vergrößert dargestellt.

Hier kann der Benutzer Ziele erstellen und hierarchische Unterziele zu gegebenen Zielen anlegen. Die Ziele werden als Baum visualisiert. Die Anzahl der Hierarchie-Ebenen ist dabei nicht eingeschränkt. Wie in der Abbildung sichtbar, sind manchen der Ziele Ressourcen zugeordnet. Der Status eines Ziels kann auf "erreicht" gesetzt werden, indem dieses mit einem Häkchen-Symbol versehen wird, d.h. Ziele werden "abgehakt". Werden Ziele abgehakt, die zu keinem Oberziel gehören, werden sie unter "Abgehakte Ziele" einsortiert und können durch Einklappen dieses Elements ausgeblendet werden. Recherchiert ein Benutzer im Web für ein aktuell von ihm verfolgtes Ziel, so kann er innerhalb seines Zielbaumes ein Ziel aktivieren. Dieses wird dann entsprechend hervorgehoben und allen nachfolgend als Ressource gespeicherten Webseiten wird das aktivierte Ziel als Tag vorgeschlagen.

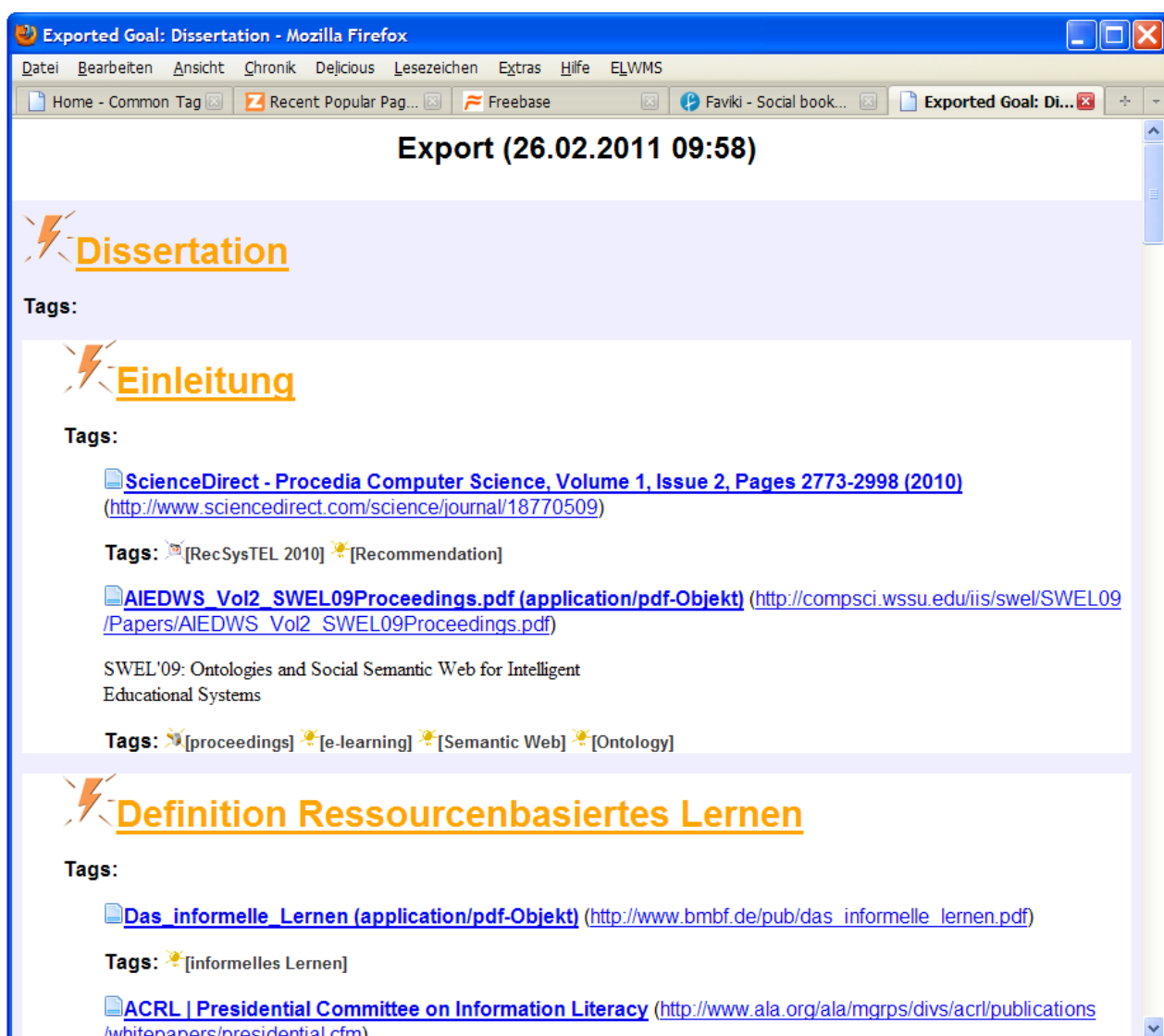


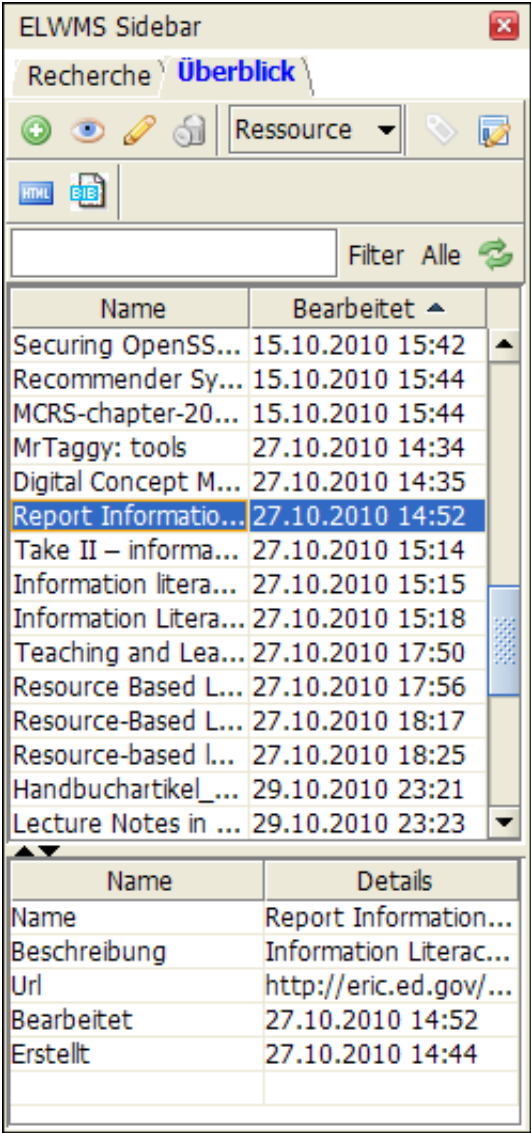
Abbildung 8.10: Export der Ziele und Ressourcen als HTML-Seite

Für eine weitere Nutzung der zu den Zielen gesammelten Ressourcen (vgl. Konzept zum Prozessschritt *Nutzen* im Modell) bzw. für eine Weitergabe an anderen Lernende (vgl. Konzept zum Prozessschritt

Weitergeben im Modell) lässt sich ein ausgewählter Teil der Ziele und deren zugeordnete Ressourcen exportieren. Bei Aufruf der Exportfunktion wird eine HTML-Seite generiert und im Webbrowser geöffnet, die die ausgewählten Ziele, die den Zielen zugeordneten Ressourcen inklusive URL, Snippets und deren Tags enthält (siehe Screenshot 8.10). Des Weiteren kann diese Ansicht den Nutzer bei der Nutzung der Ressourcen unterstützen.

Suchen und Überblick über Ressourcen im Wissensnetz

Im Tabulator **Überblick** sind verschiedene Funktionen angeordnet, die dem Benutzer unmittelbar im Add-on zur Verfügung stehen um einen Überblick über die von ihm recherchierten und dem Wissensnetz hinzugefügten Webseiten zu gewinnen und nach Ressourcen zu suchen. Diese Funktionen findet der Benutzer alternativ auch im nachfolgend beschriebenen Knowledge-Portal.



ELWMS Sidebar

Recherche Überblick

Recherche: [Suchfeld] Filter Alle

Name	Bearbeitet
Securing OpenSS...	15.10.2010 15:42
Recommender Sy...	15.10.2010 15:44
MCRS-chapter-20...	15.10.2010 15:44
MrTaggy: tools	27.10.2010 14:34
Digital Concept M...	27.10.2010 14:35
Report Informatio...	27.10.2010 14:52
Take II – informa...	27.10.2010 15:14
Information litera...	27.10.2010 15:15
Information Litera...	27.10.2010 15:18
Teaching and Lea...	27.10.2010 17:50
Resource Based L...	27.10.2010 17:56
Resource-Based L...	27.10.2010 18:17
Resource-based l...	27.10.2010 18:25
Handbuchartikel_...	29.10.2010 23:21
Lecture Notes in ...	29.10.2010 23:23

Name	Details
Name	Report Information...
Beschreibung	Information Literac...
Url	http://eric.ed.gov/...
Bearbeitet	27.10.2010 14:52
Erstellt	27.10.2010 14:44

Abbildung 8.11: Sidebar "Überblick"

Wie in Abbildung 8.11 gezeigt, werden hier alle Ressourcen in tabellarischer Form angezeigt. Dabei ist eine Sortierung nach verschiedenen Attributen möglich, wie das Bearbeitungsdatum. So kann der Benutzer schnell einen Überblick über die Ressourcen, beispielsweise die zuletzt hinzugefügten, gewinnen. Des Weiteren stehen auch hier die Bearbeitungsfunktionen wie Erstellen, Editieren und Löschen zur Verfügung. Eine weitere Funktion innerhalb des Tabulators ist die Suche nach Ressourcen anhand ihres Namens bzw. der Anfangsbuchstaben. Zusätzlich lassen sich die Ressourcen im HTML-Format oder BibTeX-Format [106] exportieren. In der Tabelle lassen sich nicht nur alle Ressourcen auflisten, sondern es ist auch ein Wechsel zu den verschiedenen Tag-Typen möglich.

8.2.2 Knowledge-Portal

Das Knowledge-Portal dient konzeptionell weniger dem Speichern von Webseiten als Ressourcen im Rahmen des Rechercheprozesses, sondern primär einer Suche und dem Navigieren in den im Wissensnetz vorhandenen Ressourcen. Es dient zudem der Erstellung und Modifikation von Profilinformationen zu den Benutzern und soll in den zukünftigen Arbeiten insbesondere auch die Community-Funktionen lokalisieren.



Abbildung 8.12: Knowledge-Portal

Der Screenshot in Abbildung 8.12 zeigt das Knowledge-Portal, das für den Einsatz in ELWMS.KOM konfiguriert wurde. Das Knowledge-Portal lässt sich unmittelbar aus der Sidebar heraus über einen Link auf eine Ressource im Webbrowser öffnen. Alternativ kann die Startseite des Knowledge-Portals über einen Button geöffnet werden. In der Mitte der generierten Webseite wird der Inhalt eines Wissensnetzknotts angezeigt, in diesem Falle ist der Wissensnetzknoten das Thema "semantisches tagging". Unter dem Namen dieses Tags sind zuerst die Attribute und deren Werte tabellarisch, darunter deren Tags aufgelistet. Des Weiteren werden die Ressourcen, die mit diesem Thema getaggt wurden, mit Name, Url, Snippet und Vorschau des Screenshots der Webseite präsentiert. In der linken Spalte des Knowledge-Portals ist zuoberst ein Link für den Zugriff auf das Nutzerprofil angeordnet. Darunter finden sich Werkzeuge zum Durchsuchen des Wissensnetzes. Dabei kann die Suche auf bestimmte Typen von Knoten, also z.B. Ressourcen oder Tags eingeschränkt werden. Darunter sind Listen mit allen Ressourcen und Tags des eingeloggten Nutzers angeordnet, die einen Schnellzugriff ermöglichen. Die Tags werden anhand ihres Tag-Typs gruppiert. Ganz unten gibt es Links zu speziellen Seiten im Knowledge-Portal.

In der rechten Spalte des Knowledge-Portals werden der Ressource zugeordnete Tags und von den Recommender-Verfahren empfohlene Ressourcen in sogenannten Kontext-Boxen angezeigt. Die Anzeige in den Kontextboxen ist abhängig vom geöffneten Wissensnetzknoten. Im abgebildeten Screenshot werden zum Tag vom Typ Thema ein ähnliches Thema des Nutzers (oberste Box) und zwei ähnliche Themen anderer Nutzer (unterste Box) vorgeschlagen. Mit diesem Thema sind nur Personentags verknüpft, die in der Box mit dem Namen "Personen" angezeigt werden. Würden noch andere Tags zu diesem Thema existieren, hätte jeder Tag-Typ seine eigene Kontext-Box.

8.3 Zusammenfassung

In diesem Kapitel wurde die Architektur von ELWMS.KOM vorgestellt. Des Weiteren wurde anhand von Screenshots die Funktionalität von ELWMS.KOM erläutert. Es wurden in dieser Arbeit grundlegende Funktionen des Konzepts implementiert, weiterführende Funktionen, insbesondere Community-Funktionen, werden im Nachfolgeprojekt "CROKODIL" umgesetzt.



9 Evaluation von ELWMS.KOM

In dieser Arbeit wurden der implementierte Software-Prototyp ELWMS.KOM, wie er in Kapitel 8 beschrieben wurde, bzw. dessen Vorgänger durch Benutzer evaluiert. Die Ergebnisse der Evaluationen wurden insbesondere genutzt, um den Software-Prototyp iterativ weiterzuentwickeln und zu verbessern.

Dieses Kapitel ist wie folgt gegliedert. Der erste Abschnitt befasst sich mit Grundlagen der Evaluation von Software, die für das Verständnis der Abschnitte, die die durchgeführten Studien beschreiben, notwendig sind. Auf Grund der Herausforderungen bei der Evaluation des Software-Prototypen (siehe Abschnitt 9.1.3) wurden im Rahmen dieser Arbeit verschiedene Benutzerstudien durchgeführt, die unterschiedliche Evaluationsziele verfolgten. Die durchgeführten Studien werden in den Abschnitten 9.2 und 9.3 näher erläutert, dabei werden bei jeder Studie die Ziele, der Ablauf und wichtige Ergebnisse beschrieben.

9.1 Grundlegendes

Ansätze und Methoden zur Evaluation von Software lassen sich im Wesentlichen im Forschungsbereich der Mensch-Computer-Interaktion (englisch Human-Computer Interaction, HCI) finden. Die in dieser Arbeit durchgeführten Studien lassen sich in Laboruntersuchungen und Felduntersuchungen einteilen. Vor- und Nachteile der beiden Ansätze werden in Abschnitt 9.1.1 beschrieben. Die im Rahmen der Arbeit angewendeten Evaluationsmethoden werden in Abschnitt 9.1.2 beschrieben. In Abschnitt 9.1.3 werden Herausforderungen erläutert, die sich aus der Evaluation des Software-Prototypen während des Einsatzes im Ressourcen-basierten Lernen ergeben.

9.1.1 Laboruntersuchung versus Felduntersuchung

Sogenannte Laboruntersuchungen stellen eine Möglichkeit der Evaluation dar, in denen die Evaluation in einer kontrollierten Umgebung durchgeführt wird. In solchen Studien werden Faktoren, die die Evaluation störend beeinflussen können, so weit wie möglich reduziert. Da für solche Studien die Probanden in der Regel in die Laborumgebung eingeladen werden müssen und die Probanden meist wenig Zeit haben, werden die Studien im allgemeinen innerhalb einer kurzen Zeitspanne durchgeführt (z.B. wenige Stunden oder Tage). Durch die Kontrolle der Störfaktoren besitzen zwar die Ergebnisse intern eine hohe Validität, da sich untersuchte Effekte mit hoher Wahrscheinlichkeit auf die untersuchte Software zurückführen lassen, allerdings bleibt häufig die Frage offen, inwieweit sich die Ergebnisse auf Situationen außerhalb der kontrollierten Umgebung übertragen lassen.

Felduntersuchungen stellen eine Möglichkeit dar, die Software unter realistischen Bedingungen zu untersuchen. Sie finden im Allgemeinen in einer durch die Evaluation möglichst unbeeinflussten Umgebung statt, beispielsweise direkt am Arbeitsplatz der Probanden. Dadurch sind die Ergebnisse realistischer als in Laboruntersuchungen. Sie haben allerdings den Nachteil, dass die untersuchten Effekte auf verschiedene Ursachen zurückgeführt werden können, da störende Einflussfaktoren nicht kontrolliert werden können. In dieser Arbeit werden beide Evaluationsformen genutzt.

9.1.2 Evaluationsmethoden

In dieser Arbeit wurden folgende Methoden zur Evaluation des Prototypen eingesetzt:

- **Fragebogen:** Fragebögen enthalten schriftlich formulierte Fragen, die von den Teilnehmern der Evaluation beantwortet werden sollen. Sie können einerseits auf Papier ausgedruckt sein oder online zur Verfügung gestellt werden. Die Erstellung von Fragebögen ist aufwändig, da die Fragen verständlich und neutral formuliert werden müssen. Zur Erleichterung der späteren Auswertung und Vergleichbarkeit der Antworten der Teilnehmer werden häufig Antworten zur Auswahl gestellt statt frei formulierte Antworten zu erlauben. Daher muss bei der Auswahl der Antworten darauf geachtet werden, dass möglichst alle relevanten Antworten vorgegeben sind. Anders als in mündlichen Interviews kann auf die Antworten der Teilnehmer nicht reagiert und nachgefragt werden, so dass eventuell wichtige Hinweise der Teilnehmer verloren gehen. Fragebögen fragen in der Regel die Selbsteinschätzung der Teilnehmer ab, daher sind die Ergebnisse nicht objektiv, sondern subjektiv und hängen von Erinnerungsvermögen, Konzentration, Gefühlen etc. der Teilnehmer ab.
- **Usability-Test:** Usability-Tests fokussieren auf die Benutzbarkeit der Software. In der Regel müssen die Teilnehmer eine Reihe von realistischen Aufgaben mit der Software durchführen. Anschließend kann die Meinung der Teilnehmer abgefragt werden. Neben der Befragung der Teilnehmer mittels Fragebogen oder Interview kann auch eine Beobachtung der Teilnehmer während der Bearbeitung der Aufgaben stattfinden. Des Weiteren kann überprüft werden, inwieweit die Aufgaben abgeschlossen werden konnten und ob das Ergebnis den Erwartungen entspricht. Wird z.B. die Aufgabe vorgegeben, dass eine Webseite mit der Software gespeichert werden soll, kann anschließend geprüft werden, ob dies erfolgreich geschehen ist.
- **Interview:** Neben schriftlichen Fragebögen eignen sich Interviews zur qualitativen Evaluation der Software. Während der Befragung kann der Interviewende Einfluss auf den Verlauf des Interviews nehmen und bei bestimmten Antworten nachfragen. Die Antworten lassen sich in der Regel meist nur mit hohem Aufwand mit quantitativen Zahlen beschreiben, so dass die Interviews im Rahmen dieser Arbeit vor allem eingesetzt wurden, um das Verhalten der Teilnehmer besser zu verstehen und wichtige Hinweise von den Teilnehmern zu erhalten.
- **Aufzeichnung (Logging):** Bei der Evaluierung von Software in kontrollierten Umgebungen können Nutzungsdaten wie Mausklicks, Bearbeitungszeit etc. automatisch gesammelt werden. Dies ist in Feldstudien teilweise technisch nicht realisierbar, weil z.B. die spezielle Aufzeichnungssoftware nicht auf dem Computer der Teilnehmer installiert ist. Aber auch wenn das Aufzeichnen der Daten in Feldstudien technisch realisierbar ist, ist das Aufzeichnen des Nutzungsverhaltens auf Grund der Privatsphäre der Teilnehmer in Felduntersuchungen nur eingeschränkt möglich. Bei der Auswertung von Log-Daten ist zu beachten, dass sie zu Fehlinterpretationen führen können. Beispielweise kann eine lange Verweildauer dahingehend interpretiert werden, dass der Teilnehmer die Webseite intensiv liest. Der Grund für die lange Verweildauer kann aber auch der sein, dass sich der Teilnehmer eine neue Tasse Kaffee geholt hat und nicht vor dem Bildschirm saß.

9.1.3 Herausforderungen bei Evaluationen im Kontext des Ressourcen-basierten Lernens

Bei der Evaluation von ELWMS.KOM gibt es auf Grund der Tatsache, dass es sich um einen Forschungsprototypen handelt, und auf Grund des Szenarios des Ressourcen-basierten Lernens verschiedene Schwierigkeiten:

- **Langzeiteinsatz:** Für die Evaluation ist ein Langzeiteinsatz der Software notwendig. Das Problem besteht darin, dass die Software regelmäßig aktiv von den Teilnehmenden genutzt werden muss, um ihre Stärken zu zeigen, und nicht nur passiv im Hintergrund ablaufen kann. Es ist zwar davon auszugehen, dass die Zielgruppe von ELWMS.KOM regelmäßig Recherchen durchführen muss, aber es ist nicht davon auszugehen, dass sie dies regelmäßig mit ELWMS.KOM tut. Denn die Teilnehmer nutzen bereits andere Systeme bei ihren Recherchen und müssen daher motiviert werden, eine Software zu nutzen, die nur eine befristete Zeit zur Verfügung steht, d.h. die langfristige Nutzbarkeit u.a. der in die Software eingegebenen Daten kann nicht gewährleistet werden. Da die Rechercheergebnisse für die Teilnehmer in der Regel wichtig sind, werden sie ELWMS.KOM entweder parallel zu ihrer herkömmlich genutzten Software verwenden oder nur relativ wenige Recherchen mit ELWMS.KOM durchführen, da sie durch die Teilnahme an der Evaluation Mehraufwand haben.
- **Benutzbarkeit der grafischen Benutzungsoberfläche versus Benutzbarkeit der Software:** In allgemeinen Usability-Tests, bei denen eine Evaluation nur während einer kurzen Zeitspanne durchgeführt wird, werden den Teilnehmern häufig fiktive Aufgaben gestellt. Die Bearbeitung solcher Aufgaben zeigt daher in der Regel nur Schwachstellen der grafischen Benutzungsoberfläche auf. Wichtig für die Evaluation von ELWMS.KOM ist aber auch die Benutzbarkeit der Software an sich, also die Nützlichkeit der konzipierten Funktionen unabhängig von ihrer tatsächlichen Implementierung. Die Evaluation von Konzepten, die in einer Software stecken, ist aber immer durch die grafische Benutzungsoberfläche beeinflusst, d.h. eine Trennung der Konzepte von der grafischen Benutzungsoberfläche ist sehr schwierig.
- **Evaluationsteilnehmer:** Für die Evaluation war es wichtig, die Teilnehmer der Evaluation vor Ort zu haben. Dies schränkt zwar die Anzahl möglicher Teilnehmer ein, aber hat gewisse Vorteile, z.B. kann bei Problemen mit der Software direkt geholfen werden, ohne teure Anfahrtswege oder lange zeitliche Verzögerungen. Während der Langzeitevaluationen können die Teilnehmer hin und wieder motiviert werden, ELWMS.KOM zu nutzen. Durch den persönlichen Kontakt können auch informelle Gespräche über die Software geführt werden. Des Weiteren sind die Teilnehmer aller Evaluationen Studenten und wissenschaftliche Mitarbeiter und entsprechen damit der Zielgruppe von ELWMS.KOM.
- **Messung von Lernerfolg:** In einer Evaluation ist es grundsätzlich sehr schwierig, den Lernerfolg, der sich durch die Benutzung einer Software einstellt, zu messen. Denn der Lernerfolg wird von vielen Faktoren wie Motivation, Vorwissen, Fähigkeiten, Konzentration etc. beeinflusst. Im Rahmen dieser Arbeit wird davon ausgegangen, dass eine Software, die den Lernprozess positiv beeinflusst wie z.B. durch Anregung zur Reflektion des Lernprozesses oder Erhöhung der Konzentration der Lernenden durch Minimierung der Ablenkungen, auch einen positiven Effekt auf den Lerner-

folg haben kann. Daher wird in den Evaluationen eher auf den Lernprozess fokussiert als auf das Lernergebnis.

9.2 Durchgeführte Studien mit Vorläufern des Prototypen

Die in diesem Abschnitt beschriebenen Studien wurden mit sehr frühen Prototypen durchgeführt, um bereits zu Beginn der Arbeit Rückmeldungen zu bestimmten Aspekten von ELWMS.KOM zu bekommen und Erfahrungen zu sammeln.

9.2.1 Studie Labor 1

Der in dieser Studie verwendete Prototyp basierte noch nicht auf der in Kapitel 8 beschriebenen Client-Server-Architektur, sondern bestand nur aus dem Add-on im Webbrowser Firefox. Das Add-on bot die Möglichkeit, Ziele zu planen und Ressourcen zu diesen Zielen als Wissensnetz zu speichern. Das Tagging war zu dem Zeitpunkt nur rudimentär ohne die Unterscheidung von Tag-Typen implementiert.

Neben der Sammlung von ersten Erfahrungen beim Einsatz des Prototypen hatte die Studie weitere Evaluationsziele:

- **Analyse des Vorgehens beim Ressourcen-basierten Lernen:** Die Evaluationsteilnehmer wurden mittels schriftlichem Fragebogen befragt, wie sie im Allgemeinen recherchieren und welcher Bedarf für eine unterstützende Software besteht.
- **Weiterentwicklung des Prototypen:** Die Beobachtung der Teilnehmer während der Nutzung der Software gab Hinweise auf Verbesserungsmöglichkeiten des Prototypen. Die Beobachtung fand zum einen als stiller Beobachter hinter den Teilnehmern statt und zum anderen wurde das Nutzerverhalten mittels einer Software aufgezeichnet.
- **Untersuchung der Zielmanagementfunktion:** Es wurde untersucht, wie die Teilnehmer mit der Möglichkeit umgehen, sich vor und während einer Recherche Ziele zu setzen.
- **Untersuchung des Lernprozesses:** Vor und nach der Studie wurden die Teilnehmer mittels schriftlichem Fragebogen zu ihren Emotionen befragt. Zusätzlich wurde mittels eines Wissens-tests erhoben, ob die Nutzung der Software einen Einfluss auf den Wissenszuwachs hat.

Die Generalisierung der Ergebnisse dieser Studie und Langzeiterfahrungen waren zum Zeitpunkt der Untersuchung noch nicht wichtig. Daher wurde diese Studie als Laboruntersuchung mit Experimental- und Kontrollgruppe geplant und durchgeführt. Diese Studie wurde gemeinsam mit dem Medieninformatiker Philipp Scholl und dem Psychologen Bastian Benz durchgeführt. Ausführliche Informationen zum Versuchsdesign und Auswertungen statistischer Kennzahlen wie die Auswertung der Emotionen und des Wissenstests können in [15] und [128] nachgelesen werden. Im Folgenden werden der Ablauf und die für diese Arbeit wichtigen Ergebnisse knapp zusammengefasst.

Evaluationsdesign

Der Prototyp wurde mit einer Gruppe von 64 Psychologie-Studierenden evaluiert. Die Studierenden mussten in der Online-Enzyklopädie Wikipedia nach Informationen zum Thema "Altertum" recherchieren. Dabei nutzten 31 Studierende in der Experimentalgruppe den Prototypen, der andere Teil der

Studierenden diene als Kontrollgruppe und durfte den Prototypen nicht nutzen. Insgesamt war die Evaluation auf zwei Stunden begrenzt, abzüglich der Zeit zum Ausfüllen der Fragebögen standen für den Rechercheprozess in der Wikipedia effektiv 45 Minuten zur Verfügung.

Vor der Aufgabe absolvierten die Studierenden einen Wissenstest mit Multiple-Choice-Fragen über das gewählte Thema. Den Studierenden wurde gleich zu Beginn der Studie mitgeteilt, dass sie nach der Recherche den gleichen Wissenstest noch einmal lösen müssen und ihre im Add-on gesammelten Ressourcen dabei nicht verwenden dürfen.

Im Wesentlichen bestand die Aufgabe der Experimentalgruppe darin, sich Lernziele anzulegen, Ressourcen zu sammeln und sie den Lernzielen zuzuordnen. Aufgrund der vorgegebenen Aufgabenstellung und der Kürze der Zeit wurden allerdings hauptsächlich themenspezifische Ziele gesetzt. Die Kontrollgruppe hatte nur die Aufgabe, in der Wikipedia zu recherchieren, um anschließend die Fragen des Wissenstests erneut zu beantworten.

Nach der Aufgabenbearbeitung führten die Versuchsteilnehmer den Wissenstest erneut aus und ebenso beantworteten sie den Fragebogen zur Erhebung ihres sonstigen Rechercheverhaltens. Bei der Auswertung der Evaluation in dieser Arbeit standen vor allem die Visualisierung des Wissensnetzes und die sonstige softwareseitige Unterstützung des Rechercheprozesses im Vordergrund.

Ergebnisse

In Relation zur kurzen Dauer der Bearbeitung der Aufgabe, die die Studierenden für die Recherche zur Verfügung hatten, wurden sehr umfangreiche und qualitativ gute Wissensnetze erstellt. Im Durchschnitt wurden 13,9 Ziele angelegt und durchschnittlich 17,37 Wissensartefakte gespeichert. Die Verteilung über die Versuchsteilnehmer kann den Tabellen 9.1 und 9.2 entnommen werden, z.B. erstellten 40% der Versuchsteilnehmer zwischen 11 und 20 Ziele.

	0-10 Ziele	11-20 Ziele	11-30 Ziele	31-40 Ziele	> 41 Ziele
Teilnehmer	40 %	40 %	13 %	7 %	0 %

Tabelle 9.1: Anzahl erstellter Ziele

	0-10 Ress.	11-20 Ress.	11-30 Ress.	31-40 Ress.	> 41 Ress.
Teilnehmer	23 %	47 %	17 %	13 %	3 %

Tabelle 9.2: Anzahl gespeicherter Ressourcen

Mittels einer persönlichen Betrachtung der entstandenen Wissensnetze konnte festgestellt werden, dass die erstellten Lernziele gut zur Aufgabenstellung passen, denn sie enthalten in den meisten Fällen die Fragen aus dem Wissenstest. Die Stellen, die die Studierenden aus den Webseiten als Snippet entnommen haben, passen ebenfalls gut zu den Lernzielen. Während der Bearbeitung tauchten kaum Fragen zur Benutzung der Software auf, obwohl zu Beginn der Aufgabenbearbeitung nur eine kurze Zeit für die Einarbeitung in die Software zur Verfügung stand. Die Experimentalgruppe schnitt im Wissenstest trotz wegen der Einarbeitung in die Software geringerer zur Verfügung stehender Recherchezeit genauso gut ab wie die Kontrollgruppe.

Software / Hilfsmittel	Anzahl Nennung
Textverarbeitungsprogramm	83%
Notizen auf Papier	33%
Lesezeichen im Webbrowser	11%
Ausdruck	8%
Webseite auf Festplatte speichern	6%
Mind-Mapping-Tool	6%
PowerPoint	3%
Excel	3%

Tabelle 9.3: Allgemein verwendete Hilfsmittel während der Recherche

In Tabelle 9.3 sind die Antworten auf die Frage nach den von den Teilnehmern allgemein verwendeten Hilfsmitteln zur Speicherung der Ergebnisse während einer Recherche, wie sie kurz zuvor stattfand, enthalten. Mehrfachnennungen waren hier aufgrund der Freitextantwort möglich. 83% der Probanden gaben an, bei Webrecherchen normalerweise die wichtigen Stellen in ein Textverarbeitungsprogramm zu kopieren. 33% nutzen Papier und Stift, um ihre Notizen festzuhalten. 11% speichern ihre Ergebnisse als Lesezeichen im Webbrowser. Nur 6% nutzen Mind-Mapping-Tools zur Organisation ihrer Ergebnisse. Weitere Antworten können der Tabelle 9.3 entnommen werden.

77% der Teilnehmer der Experimentalgruppe gaben an, dass sie die Firefox-Erweiterung auch zukünftig gern verwenden würden.

9.2.2 Studie Seminar 1

Der Prototyp wurde nach der ersten Laborevaluation weiterentwickelt. Die Wissensnetze im weiterentwickelten Prototypen, wie er in dieser Studie verwendet wurde, werden nicht mehr lokal gespeichert, sondern die Client-Server-Architektur (siehe Implementierungskapitel 8) wurde umgesetzt. Des Weiteren wurden die typisierten Tags eingeführt. Bis auf wenige Funktionen wie z.B. manche der Exportfunktionen und Teile der Recommendation-Verfahren waren bereits alle Funktionen integriert, die im Implementierungs-Kapitel beschrieben sind.

Da die Ergebnisse einer Langzeitnutzung für die Weiterentwicklung des Prototypen immer wichtiger wurden, wurde eine Felduntersuchung geplant. Die Evaluation hatte folgende Ziele:

- **Weiterentwicklung des Prototypen:** Zum einen spielt hier der Aspekt der Benutzbarkeit eine Rolle, d.h. ob die Teilnehmer der Evaluation in der Lage sind, den weiterentwickelten, komplexeren Prototypen zu nutzen. Während der Nutzung des Prototypen wurden Mausklicks auf die Buttons des Prototypen aufgezeichnet, um herauszufinden, ob es Funktionen gibt, die von den Teilnehmern nie genutzt werden. Zusätzlich wurden die Teilnehmer qualitativ befragt, welche Änderungen an der Software sie sich wünschen würden.
- **Untersuchung der Verwendung der typisierten Tags:** Durch die Auswertung der Wissensnetze wurden Erkenntnisse zur Nutzung der typisierten Tags gewonnen. Die Auswertung wurde bereits in Kapitel 4.3.2 beschrieben.

Die Evaluation fand mit 12 Studierenden des Seminars "Advanced Research Topics in E-Learning Technology" im Sommersemester 2009 statt. Zu Beginn des Seminars wurde der Prototyp kurz demonstriert. Die Studierenden hatten dann die Aufgabe, den Prototypen zuhause auf ihrem Computer zu installieren und sich damit vertraut zu machen. Im Seminar musste eine Hausarbeit angefertigt werden, für die auch selbstständig recherchiert werden musste. Während der Recherche sollten die Studierenden die Software regelmäßig nutzen und ein Wissensnetz erstellen. Am Ende des Seminars wurde der Inhalt der Hausarbeit in einem Vortrag vorgestellt, zusätzlich präsentierten die Studierenden ihre Wissensnetze und es fand eine Diskussion über ihre Erfahrungen mit der Software statt.

Während der Recherchephase, d.h. ca. 4 Wochen nach Beginn der Nutzung des Prototypen, wurde die Benutzungsoberfläche des Prototypen leicht verändert, um zu sehen, ob das Verschieben der Zielmanagementfunktion aus dem Blickfeld der Teilnehmer Auswirkungen auf das Nutzungsverhalten hat. Zu Beginn wurde der Zielbaum im vordersten Tab des Prototypen angezeigt, später wurde er auf den zweiten Tab verschoben.

An jedem Nutzungstag öffnete sich automatisch ein kurzer Online-Fragebogen im Webbrowser, der von den Studierenden ausgefüllt werden sollte.

Ergebnisse

Das Seminar eignete sich gut zur Untersuchung der Software über einen längeren Zeitraum. Denn das Thema des Seminars war von den restlichen Inhalten des Studiums der Teilnehmer stark abgegrenzt, so dass alle Inhalte für die Hausarbeit neu recherchiert werden mussten.

Von den Studierenden wurden nicht alle möglichen Funktionen der Software verwendet. Zwar fand teilweise ein Testen der Funktionen zu Beginn der Evaluation statt, gegen Ende hin wurden aber nur noch Speicher- und Zielmanagementfunktionen genutzt.

Das Verschieben der Zielmanagementfunktionen auf den zweiten Tab hatte keine Auswirkungen auf das Nutzerverhalten. Die Zielmanagementfunktionen wurden dennoch verwendet. Es ist daher davon auszugehen, dass die Zielmanagementfunktionen wichtig sind und die Planung der Recherche auch fortgesetzt wird, wenn die Zielmanagementfunktion aus dem Blickfeld der Teilnehmer verschwindet.

In Diskussionen mit den Teilnehmern wurde festgestellt, dass nur zwei der Teilnehmer überhaupt vor Beginn der Studie Tagging kannten und bisher noch keiner der Teilnehmer selbst etwas getaggt hatte. Daher wurde in den folgenden Studien der Prototyp nicht nur demonstriert und insbesondere auf die Funktion des Taggings eingegangen, sondern es wurde auch ein Lernvideo erstellt mit kurzen Aufgaben, damit sich die Teilnehmer leichter in die Funktionen des Prototypen einarbeiten können.

Der Online-Fragebogen, der von den Teilnehmern täglich nach jeder Benutzung der Software ausgefüllt werden sollte, wurde sehr selten bis nie ausgefüllt. Dies mag zum einen daran liegen, dass die Teilnehmer generell an nur wenigen Tagen für ihre Seminararbeit recherchiert haben, und zum anderen daran, dass der Zeitpunkt zum automatischen Öffnen des Fragebogens scheinbar schlecht gewählt war. Entweder die Teilnehmer hatten an diesem Tag vorher noch keine Recherchen durchgeführt oder sie haben an diesem Tag ihren Computer gar nicht benutzt. Manuell hat keiner der Teilnehmer den Frage-

bogen geöffnet. Für die folgenden Studien wurde hieraus gelernt, dass die Rücklaufquote von täglichen Fragebögen sehr gering ist und in Zukunft darauf verzichtet wird.

Während der Nutzung stellten die Teilnehmer kleine Implementierungsfehler fest, dennoch konnte jeder von ihnen ein Wissensnetz bei ihrer Recherche erstellen. Während der finalen Feedback-Diskussionen teilten uns die Teilnehmer mit, dass sie das Werkzeug für ihre Recherchen nützlich fanden und es weiterhin nutzen würden, wenn es über den Status eines Forschungsprototypen hinaus ist.

9.3 Durchgeführte Studien mit ELWMS.KOM

Der Prototyp wurde weiterentwickelt und die bei der vorangegangenen Studie festgestellten Implementierungsfehler wurden behoben. Des Weiteren wurden die weiter entwickelten Recommender-Verfahren integriert. Der evaluierte Prototyp in den in diesem Abschnitt beschriebenen Studien ist auf dem Stand des Implementierungs-Kapitels.

9.3.1 Studie Labor 2

Die Studie hatte folgendes Evaluationsziel:

- **Untersuchung des Lernprozesses:** Es wurde untersucht, ob die Software einen Einfluss auf den Lernprozess hat.

Da die Einflussfaktoren auf das Lernen sehr groß sind, wurde entschieden, diese Studie als Laboruntersuchung mit Kontroll- und Experimentalgruppe durchzuführen.

Evaluationsdesign

An dieser Studie nahmen 8 Studierende teil, darunter eine Frau. Der Versuch dauerte insgesamt 2h, wobei ein Teil der Zeit für das Ausfüllen der Fragebogen zu Beginn und am Ende verwendet wurde, so dass insgesamt 90 Minuten für die Recherche zur Verfügung standen. Die Studierenden mussten sich vor dem Versuch zuhause mit der Software und einem Videotutorial vertraut machen, was alle auch getan haben. Jeder der Studierenden war in der Lage, die Aufgaben des Videotutorials erfolgreich zu erledigen. Die Aufgaben im Videotutorial bestanden u.a. aus der Installation der Software, Erstellen und Bearbeiten erster Ziele und Sammeln von Webseiten. Vor Beginn des Versuchs füllten die Studierenden einen Fragebogen zu demographischen Angaben und Aspekten ihres allgemeinen Rechercheverhaltens aus. Das Ausfüllen dieses Fragebogens fand noch vor der ersten Benutzung der Software statt. Danach sollten die Evaluationsteilnehmer zum Thema "Standards im E-Learning" recherchieren. Eine Gruppe, die Experimentalgruppe, bestehend aus 4 Teilnehmern sollte dafür ELWMS.KOM nutzen. Die Ergebnisse ihrer Recherche sollten die Teilnehmer in einem Textdokument zusammenfassen und am Ende abgeben. In ELWMS.KOM war bereits ein Wissensnetz zum Thema aufgebaut, das eine Themenstruktur und Ressourcen zu den Themen enthielt. Die Kontrollgruppe, die nicht mit dem Prototypen gearbeitet hat, bekam ein Placebo-Add-on im Webbrowser, mittels dessen sich eine Linkliste öffnete, die alle Ressourcen aus dem Wissensnetz alphabetisch sortiert enthielt, um ähnliche Chancen zwischen den Gruppen herzustellen. Nach der Recherche mussten die Teilnehmer erneut einen Fragebogen ausfüllen. Alle Fragebögen

wurden auf Papier ausgedruckt, da das Ausfüllen des ersten Fragebogens in einem Raum ohne Computer stattfand. Problematisch an Papierfragebögen ist, dass Fragen ausgelassen werden können, dies ist aber insgesamt nur bei zwei Fragen der Fall gewesen. Durch das Ausfüllen vor Ort konnte vermieden werden, dass die Teilnehmer den Fragebogen zu unterschiedlichen Zeiten ausfüllen oder es gar vergessen.

Ergebnisse

Die Teilnehmer der Experimentalgruppe haben kein einziges Mal das bereits existierende Wissensnetz aufgerufen, während die Teilnehmer der Kontrollgruppe die Linkliste von oben nach unten durchgeklickt haben. Die abgegebenen Textdokumente waren dennoch von gleicher Qualität. Sie wurden manuell auf korrekte Fakten und ihre Struktur hin geprüft. Demzufolge ließ sich kein Unterschied am Ergebnis der Recherche feststellen.

Allerdings konnten Unterschiede zwischen den Gruppen bezogen auf den Lernprozess festgestellt werden. Aufgrund der geringen Teilnehmerzahl soll an dieser Stelle verzichtet werden, signifikante Werte zu berechnen, da diese Daten nur Trends zeigen. Für aussagekräftigere Werte müsste die Evaluation mit einer Vielzahl an Teilnehmern wiederholt werden.

In Tabelle 9.4 werden die Mittelwerte der Antworten von Experimental- und Kontrollgruppe nach der durchgeführten Recherche gegenübergestellt. Ein hoher Mittelwert bedeutet eine hohe Zustimmung, ein geringer Wert eine geringe Zustimmung zur Frage. Im Anhang befinden sich Tabellen zur genaueren Aufschlüsselung der Verteilung der Antworten (Tabelle B.4 unabhängig von der Gruppe, Tabelle B.5 für Experimentalgruppe und Tabelle B.6 für Kontrollgruppe). Die vierte Spalte gibt an, ob sich ELWMS.KOM positiv oder negativ auf den Lernprozess auswirkt, dabei steht ein "+" für einen positiven Effekt. Die letzte Spalte gibt die Differenz der beiden Mittelwerte an. Je höher diese Differenz, desto stärker ist der Unterschied zwischen beiden Gruppen.

Beispielsweise gibt es einen sehr hohen, positiven Effekt bei der Frage, ob sich die Teilnehmer bei der Recherche bewusst gemacht haben, wie sie voran kommen und was sie noch machen müssen (Zeile 1).

Die Experimentalgruppe hat sich etwas gestresster gefühlt als die Kontrollgruppe. Dies mag daran gelegen haben, dass die Experimentalgruppe durch die Nutzung des Prototypen einen Mehraufwand zu bewältigen hatte. Sie mussten ihre Ressourcen in ELWMS.KOM abspeichern, während die Kontrollgruppe die wichtigen Abschnitte der Webseiten direkt in das Textdokument übernehmen konnte.

Auch die relevanten Antworten des vor der Recherche ausgefüllten Fragebogens befinden sich im Anhang (siehe Tabelle B.1, B.2 und B.3, wiederum aufgeschlüsselt in die beiden Gruppen). Viele der aufgelisteten Unterschiede sind unabhängig von Vorwissen oder Fähigkeiten der Teilnehmer. Teilweise allerdings könnten die Unterschiede der nach der Benutzung der Software gegebenen Antworten auf die Unterschiede der beiden Gruppen zurückzuführen sein. Beispielsweise könnte angenommen werden, dass die Experimentalgruppe bereits vor der Nutzung der Software ihre Ziele im Allgemeinen vorher plant. Tabelle 9.5 stellt die Antworten der beiden Gruppen vor der Recherche und nach der Recherche gegenüber. Der Mittelwert der Experimentalgruppe sagt aus, dass sie eher seltener vor einer Recherche planen als die Teilnehmer der Kontrollgruppe. Nach der Recherche gaben die Teilnehmer der Experimentalgruppe an, dass sie vor der Recherche geplant haben, während der Mittelwert der Kontrollgruppe sogar abnahm, so dass der Gruppenunterschied unter Berücksichtigung der Antworten vor der Recherche noch deutlicher wird.

	Experimentalgruppe	Kontrollgruppe	Effekt	Unterschied
Während der Recherche habe ich mir regelmäßig bewusst gemacht, wie ich voran komme und was ich noch machen muss.	5,75	4,25	+	1,50
Während der Recherche habe ich mir Mühe gegeben.	5,25	4,50	+	0,75
Während der Bearbeitung der Aufgabe war ich genervt.	3,00	2,75	-	0,25
Ich habe mich angestrengt, weil ich in der Aufgabe gut abschneiden wollte.	5,50	4,00	+	1,50
Ich habe konzentriert gearbeitet.	5,50	3,50	+	2,00
Ich fand es schwierig, Zusammenhänge zwischen den Informationen auf den unterschiedlichen Webseiten herzustellen.	3,00	3,25	+	-0,25
Während der Aufgabe habe ich mich gestresst gefühlt.	3,75	3,00	-	0,75
Mit meiner Arbeitsweise bin ich zufrieden.	4,25	4,00	+	0,25
Während der Recherche habe ich mir Sorgen gemacht, ob ich die Aufgabe schaffe.	3,75	3,75	o	0,00
Während der Recherche habe ich mich wohl gefühlt.	4,25	4,00	+	0,25
Ich bin mit dem Ergebnis dieser Aufgabe zufrieden.	3,75	3,25	+	0,50
Während der Recherche hatte ich Probleme, das Gelesene mit dem zu verbinden, was ich schon wusste.	2,25	2,50	+	-0,25
Die Bearbeitung der Aufgabe hat mir Spaß gemacht.	4,50	3,50	+	1,00
Die Rechercheaufgabe hat mich überfordert.	2,50	2,50	o	0,00
Die Bearbeitungszeit für diese Aufgabe war zu kurz.	3,25	3,00	-	0,25

Tabelle 9.4: Unterschiede zwischen Experimental- und Kontrollgruppe in Labor 2

	Aussage vor der Recherche: Ich plane vor der Recherche, wie ich vorgehen möchte.	Aussage nach der Recherche: Vor der Recherche habe ich mir überlegt, wie ich für diese Aufgabe vorgehen möchte.
Mittelwert der Experimentalgruppe	3,50	5,00
Mittelwert der Kontrollgruppe	4,00	3,25

Tabelle 9.5: Unterschiede des Zielmanagements zwischen Experimental- und Kontrollgruppe

9.3.2 Studie Seminar 2

Die Studie hatte folgende Evaluationsziele:

- **Untersuchung der Verwendung der typisierten Tags:** Durch die Auswertung der Wissensnetze wurden Erkenntnisse zur Nutzung der typisierten Tags gewonnen. Die Auswertung wurde bereits in Kapitel 4.3.2 beschrieben.
- **Untersuchung des Lernprozesses:** Es wurde untersucht, ob die Software durch eine längere Nutzungsdauer einen Einfluss auf den Recherche- und Lernprozess hat.

Evaluationsdesign

An dieser Studie nahmen 7 der 8 Studierenden aus Studie "Labor 2" teil, die das Seminar "Advanced Research Topics in E-Learning Technology" im Sommersemester 2010 besuchten. Das Evaluationsdesign entspricht dem aus der Studie "Seminar 1". Die Studierenden waren durch das Lernvideo bereits in den Prototypen eingearbeitet und sollten ihn während der Anfertigung ihrer Seminararbeit nutzen. Auf die täglichen Fragebögen des ersten Seminars wurde verzichtet, es wurde nur am Ende des Seminars ein Fragebogen zur Einschätzung des Prototypen durch die Teilnehmer ausgeteilt.

Ergebnisse

Tabelle 9.6 listet die Mittelwerte der Antworten im Fragebogen vor und nach der Nutzung von ELWMS.KOM auf.

	Vorher	Nachher	Effekt	Unterschied
Ich habe Mühe, Literatur für meine Vorträge und Arbeiten zu finden.	3,25	4,43	-	1,18
Ich plane vor der Recherche, wie ich vorgehen möchte.	3,75	4,86	+	1,11
Ich kann die Qualität und Vertrauenswürdigkeit der gefundenen Webseiten einschätzen.	4,25	4,29	+	0,04
Ich habe Probleme beim Wiederfinden von Seiten, die ich zu einem früheren Zeitpunkt besucht habe.	2,63	2,71	-	0,08
Ich weiß nicht, wie ich weitersuchen kann, um bessere Suchergebnisse zu bekommen.	3,38	3,57	-	0,19
Ich finde bei der Suche im Internet, wonach ich suche.	4,50	4,86	+	0,36
Mir fällt es schwer zu beurteilen, ob eine Webseite relevant für meine Recherche ist.	2,25	2,14	-	-0,11
Ich kann die gefundenen Informationen thematisch einordnen.	5,13	4,86	-	-0,27
Mich überfordert die große Menge an Informationen im Internet.	2,75	2,86	-	0,11

Tabelle 9.6: Unterschiede im Seminar vor Nutzung von ELWMS.KOM und am Ende des Seminars

Zwischen den Aussagen zum Rechercheverhalten vor der Nutzung von ELWMS.KOM und nach Ende des Seminars konnten bis auf wenige Aussagen nur geringe Unterschiede festgestellt werden. Eine Zahl mit positivem Vorzeichen in der dritten Spalte der Tabelle deutet auf einen eher positiven Effekt hin. Bei der Frage nach der Planung vor den Recherchen konnte wie in der Studie "Labor 2" ein deutlich positiver Unterschied festgestellt werden, d.h. die Studierenden gaben an, dass sie für die Recherchen eher geplant haben als normalerweise. Sie gaben allerdings an, dass sie mit der Suche nach Literatur mehr Mühe hatten. Die Vermutung liegt nahe, dass das von den sonstigen Themen im Studium weit entfernte Thema mehr Mühe bereitet hat, da sie in dem Thema kaum Vorwissen hatten. Positiv am Einsatz des Werkzeugs im Seminar war, dass die fachlichen Betreuer der Seminararbeiten Einblick in das Wissensnetz nehmen konnten und demzufolge Hinweise zur Recherche geben konnten. Eine weitere Tabelle zur Einschätzung und Verteilung der Antworten befindet sich im Anhang (siehe Tabelle B.7).

Diese in Tabelle 9.6 gezeigten Fragen, die das Rechercheverhalten vorher und nachher gegenüberstellen, eignen sich bei geringer Teilnehmerzahl eher wenig, um eventuelle Effekte der Software zu zeigen, wenn die Effekte eher gering sind. Auf der 6-stufigen-Antwortskala von "Stimme voll und ganz zu" bis "Stimme überhaupt nicht zu" muss der Effekt relativ groß sein, um die Meinung der Teilnehmer von z.B. "Stimme überwiegend zu" in "Stimme voll und ganz zu" zu ändern. Eine Ausdehnung auf eine höherstufige Skala ist meist nicht praktikabel, da die Interpretation der Antworten zeitlich und zwischen den Teilnehmern eher schwanken kann, und die Antworten der Teilnehmer somit weniger vergleichbar werden.

Daher wurden den Teilnehmern auch Fragen gestellt, die sich konkret auf den Prototypen beziehen, z.B. ob die Software die Anzahl der Wechsel zwischen Webbrowser und anderen Anwendungen reduzieren konnte. In Tabelle B.7 im Anhang sind diese Fragen und die Antworten der Teilnehmer aufgelistet.

In Tabelle 9.7 wird die Gesamt-Einschätzung der Teilnehmer über die Software am Ende des Seminars angegeben. Die Teilnehmer haben ELWMS.KOM eher gern genutzt und stimmten eher zu, dass ELWMS.KOM auf die Anforderungen der Rechercheaufgabe zugeschnitten war. Teilweise gingen die

	Stimme voll und ganz zu (6)	Stimme über- wie- gend zu (5)	Stimme eher zu (4)	Stimme eher nicht zu (3)	Stimme über- wie- gend nicht zu (2)	Stimme über- haupt nicht zu (1)	Mittelwert	Standard- abweichung
Ich habe ELWMS für meine Recherchen während meiner Seminararbeit gern genutzt.	-	3	2	2	-	-	4,14	0,90
ELWMS war sehr gut auf die Anforderungen der Rechercheaufgabe zugeschnitten.	1	-	5	1	-	-	4,14	0,90
ELWMS hat mir geholfen, bei der Recherche strukturierter vorzugehen als normalerweise.	-	1	3	2	1	-	3,57	0,98
Ich würde ELWMS meinen Freunden empfehlen, wenn diese Recherchen durchführen und ihre Webseiten verwalten müssen.	2	3	2	-	-	-	4,29	1,25
ELWMS war mir bei der Organisation der gefundenen Quellen (Webseiten) nützlich	-	2	4	1	-	-	4,14	0,69
Die Qualität meiner Seminararbeit schätze ich durch die Nutzung von ELWMS besser ein.	-	1	3	3	-	-	3,71	0,76
Ich würde ELWMS gern für weitere Recherchen nutzen.	1	1	2	1	2	-	3,71	1,50

Tabelle 9.7: Meinung der Teilnehmer über ELWMS.KOM nach Ende des Seminars

Meinungen aber stark auseinander, z.B. bei der Frage, ob sie ELWMS.KOM weiterhin nutzen würden. Allerdings würde jeder der Teilnehmer die Software seinen Freunden weiterempfehlen. Insgesamt war die Einschätzung des Prototypen eher positiv.

9.3.3 Studie KOM

Die Studie hatte folgende Evaluationsziele:

- **Analyse des Vorgehens beim Ressourcen-basierten Lernen:** Die Evaluationsteilnehmer wurden mittels schriftlichem Fragebogen befragt, wie sie im Allgemeinen recherchieren und welcher Bedarf für eine unterstützende Software besteht.
- **Untersuchung der Verwendung der typisierten Tags:** Durch die Auswertung der Wissensnetze wurden Erkenntnisse zur Nutzung der typisierten Tags gewonnen. Die Auswertung wurde bereits in Kapitel 4.3.2 beschrieben.
- **Untersuchung der Benutzbarkeit von ELWMS.KOM:** Es wurde untersucht, ob die Software während einer längeren Nutzungsdauer außerhalb eines Seminars Schwächen bei der Benutzbarkeit aufweist.
- **Untersuchung des Lernprozesses:** Es wurde untersucht, ob die Software bei einer längeren Nutzungsdauer einen Einfluss auf den Recherche- und Lernprozess hat. Hierbei sollte gezeigt werden, ob die Software positive Effekte auf die strukturelle bzw. konzeptuelle Desorientierung und die durch den Aufwand des Ressourcenmanagements entstehende, kognitive Mehrbelastung hat.

Evaluationsdesign

An der Befragung nahmen 36 Universitäts-Mitarbeiter und 17 Studierende teil, darunter 8 Frauen. Die Studierenden waren dabei nicht Teilnehmer eines Seminars. Der erste Teil der Studie bestand aus dem

Befragen der Teilnehmer mittels eines Online-Fragebogens nach ihrem allgemeinen Rechercheverhalten und ihrem Bedarf. Der Abschnitt B.4 im Anhang enthält den Fragebogen. Die Antworten des Fragebogens zum Rechercheverhalten sind in Tabelle B.9 im Anhang enthalten. Die wichtigsten Ergebnisse der Befragung wurden in Kapitel 2 bereits erläutert.

Einer Langzeitnutzung von ELWMS.KOM stimmten 20 Teilnehmer zu. Teilweise war es den anderen Befragten nicht möglich, an der Studie teilzunehmen, weil der Prototyp aufgrund der Software-Abhängigkeiten z.B. nicht auf einem Mac installiert werden kann. Die Teilnehmer arbeiteten sich in ELWMS.KOM mittels des zur Verfügung gestellten Lernvideos ein und sollten den Prototypen möglichst intensiv nutzen. Nach einigen Wochen erhielten die Teilnehmer eine Mail mit dem Link zum zweiten Fragebogen, der sie zur ihrer Einschätzung von ELWMS.KOM befragte.

In der Studie machten sich die in Abschnitt 9.1.3 genannten Probleme im Zusammenhang mit einer Langzeitstudie stark bemerkbar. Der Prototyp ist von der Nutzung des Webbrowsers Firefox und von einer Java-Version abhängig. Im Verlauf der Studie gab es Updates des Webbrowsers, wodurch Java-Applets nicht mehr im Webbrowser gestartet werden konnten und deshalb auch der Prototyp kurzfristig nicht genutzt werden konnte, bis auch für Java Updates zur Verfügung gestellt wurden. Daher hatten die Nutzer doppelten Aufwand, wenn sie wichtige Recherchen sowohl mit ELWMS.KOM als auch ihren Standard-Anwendungen durchgeführt haben. Manche Teilnehmer teilten mit, dass sie in der Evaluationszeit nur wenige Recherchen durchführen mussten und ELWMS.KOM deshalb eher selten genutzt haben. Nichtsdestotrotz haben alle Teilnehmer der Evaluation versucht, ihren Eindruck von ELWMS.KOM beim Ausfüllen des Fragebogens wiederzugeben. Der Fragebogen ist im Anhang in Abschnitt B.6 enthalten.

Ergebnisse

In einem Teil des Fragebogens wurden die Teilnehmer zur Usability befragt. Da es sich bei ELWMS.KOM um einen Forschungsprototypen handelt, wurde verzichtet einen Standard-Fragebogen zur Ergonomie-Norm DIN EN ISO 9241 (Ergonomie der Mensch-System-Interaktion) zu verwenden, wie er beispielsweise auf Webseiten zum Thema Usability¹ zu finden ist. Die meisten Fragen dieses Fragebogens sind sehr allgemein gehalten und zielen auf die Effizienz von Software-Systemen ab. Daher wurde in dieser Arbeit der Fragebogen zur Ermittlung eines Usability-Scores auf der "System Usability Scale (SUS)" [22] verwendet. Mit wenigen Fragen wird ein Score berechnet, der zwischen 0 und 100 liegt, wobei 100 der beste Wert ist.

Nutzungshäufigkeit	Mittelwert	Standardabweichung
Alle	68,63	17,23
Nie	45,0	(nur 1 Person)
Selten	70,7	20,75
Ab und zu	67,9	18,56
Sehr häufig	71,5	9,45

Tabelle 9.8: Usability (SUS Score)

¹ ErgoNorm-Benutzerfragebogen zu "Arbeit & Software", <http://www.fit-fuer-usability.de/downloads/Leitfaden-Usability.pdf>, Zuletzt abgerufen am 28.02.2011

In Tabelle 9.8 sind die ermittelten Usability-Scores insgesamt und für bestimmte Nutzungshäufigkeiten angegeben. Je häufiger das System genutzt wurde, desto höher ist unter Einbeziehung der Standardabweichung der Usability-Score. Daraus lässt sich schließen, dass das System eine gewisse Einarbeitungszeit benötigt. Sicherlich kann auch die Benutzungsoberfläche noch weiter optimiert werden.

Die Teilnehmer wurden befragt, ob sie im Prototypen überflüssige Eingaben machen mussten (siehe Tabelle 9.9). 50% der Teilnehmer gaben an, dass sie keine überflüssigen Eingaben machen mussten, 40% wenige und nur 10% einige überflüssige Eingaben.

	Einige überflüssige Eingaben	Wenige überflüssige Eingaben	Keine überflüssigen Eingaben
Hatten Sie das Gefühl, bei der Arbeit mit ELWMS Eingaben gemacht zu haben, welche Ihnen eigentlich überflüssig erschienen?	10%	40%	50%

Tabelle 9.9: Einschätzung überflüssiger Eingaben in ELWMS.KOM

Als überflüssige Eingaben wurden die folgenden Antworten genannt:

- manuelles Kopieren aus PDFs,
- Verwendung von Tags, die später nicht mehr verwendet wurden,
- einige Mausklicks, da es keine Funktionszugriffe über Tastatureingaben ("Hotkeys") gibt, und
- beim Kopieren von mehreren Snippets aus der gleichen Webseite musste der Tagging-Dialog erneut ausgefüllt werden.

Die Antwort, dass eingegebene Tags später nicht mehr verwendet wurden, wurde von einigen der Teilnehmer gegeben. Im Konzept sind zwar Pflegefunktionen des Wissensnetzes vorgesehen, diese sind aber bisher nicht im Prototypen implementiert. Eine solche Funktionalität könnte den Nutzern jedoch dabei helfen, das Wissensnetz "aufzuräumen" und selten verwendete Tags zu identifizieren und ggf. durch andere Tags zu ersetzen.

Die Teilnehmer der Evaluation wurden des Weiteren befragt, ob sie sich in ELWMS.KOM weitere Funktionen wünschen würden. 60% verneinten diese Antwort. Einige der bejahenden Teilnehmer nannten folgende Funktionswünsche:

- In die Snippets sollten auch Bilder aus den Webseiten gespeichert werden können. Momentan ist nur das Übernehmen von Text aus Webseiten möglich.
- Zusätzlich zu den manuellen Tags sollten Tags auch maschinell erstellt werden.
- Aktuell wird nur das Kopieren aus textbasierte Webseiten in die Snippets unterstützt, viele der genutzten Ressourcen sind aber PDF-Dokumente und mussten daher manuell kopiert werden. Deshalb gab es den Wunsch, auch PDF-Dokumente zu unterstützen.
- "Piling & Filing": Es besteht der Wunsch nach einer Art "Schuhkiste", in die schnell Webseiten unstrukturiert gesammelt werden können, die in einem späteren Arbeitsschritt erst strukturiert oder aussortiert werden.
- Momentan ist es nur möglich, genau einen Link für eine Ressource anzugeben. Daher bestand ein Wunsch aus der gleichzeitigen Angabe mehrerer Links.
- Es wurde sich für das Zielmanagement ein Zeitbalken gewünscht, der die verbliebene Zeit bis zu einer manuell gesetzten Deadline anzeigt.

- Aufgaben/Ziele sollten mit anderen Nutzern von ELWMS.KOM geteilt werden können, so dass verschiedene Nutzer gemeinsam ein Ziel bearbeiten können. Dies ist im Konzept vorgesehen, aber noch nicht in ELWMS.KOM integriert.
- Es wurde der Wunsch genannt, eine direkte Anbindung an bestehende Literatur/ Recherchedatenbanken zur Verfügung zu stellen.

Die Wünsche der Teilnehmer sollten bei einer Weiterentwicklung des Prototypen berücksichtigt werden, teilweise sind sie auch bereits im Konzept integriert (z.B. die Community-Funktionen).

Die Teilnehmer wurden des Weiteren befragt, wie wichtig ihnen bestimmte Funktionen in ELWMS.KOM sind. Die Antworten sind in Tabelle 9.10 aufgelistet. Manche der Teilnehmer wussten bei einigen der Funktionen nicht, was damit gemeint ist (siehe letzte Spalte). Zu den Funktionen, die hier einen hohen Anteil haben, sollten die zukünftige Benutzeroberfläche optimiert und die Hilfedokumentation entsprechend angepasst werden. Die meisten der Funktionen von ELWMS.KOM werden als unbedingt notwendig (erste Spalte) oder wenigstens ab und zu sinnvoll (zweite Spalte) eingestuft. Nur für die Möglichkeit, sich eigene Erweiterungen zu definieren, sieht eine größere Teilnehmerzahl keinen Bedarf (dritte Spalte). Eventuell ergibt sich der Bedarf bei wenigen Nutzern gar nicht oder er wird erst bei einer sehr viel längeren Nutzungsdauer bemerkbar.

Sie konnten verschiedene Funktionen in ELWMS kennenlernen und nutzen. Kreuzen Sie an, wie wichtig und sinnvoll Sie diese Funktionen für Ihre Recherchen bewerten.	Unbedingt notwendig.	Ab und zu sinnvoll.	Kein Bedarf.	Ich weiß nicht, welche Funktion gemeint ist.
Speichern von markierten/ausgewählten Abschnitten	60%	30%	0%	10%
Taggen von Webseiten	80%	20%	0%	0%
Tags, die beim Tagging vorgeschlagen werden	55%	40%	5%	0%
Angabe des Typs des Tags (z.B. Thema, Person,...)	60%	15%	20%	5%
Taggen von Tags (Verknüpfung von Tags untereinander, z.B. Person - Thema)	20%	55%	10%	15%
Aufbau einer Themenstruktur (Unterthema/Oberthema)	50%	40%	5%	5%
Möglichkeit, sich Ziele zu setzen	55%	30%	5%	10%
Zuordnung von Webseiten zu Zielen (Webseiten mit Zielen taggen bzw. Ziel aktivieren Button mit lila Punkt, um Webseiten automatisch zuzuordnen)	55%	15%	10%	20%
Anzeige in der Sidebar, ob eine Webseite bereits gespeichert ist	80%	15%	0%	5%
Weitere Webseiten werden in der Sidebar vorgeschlagen (z.B. Webseiten, die zu einem ähnlichen Thema abgespeichert wurden)	25%	55%	5%	15%
Ansicht von Webseiten und Tags im Portal (Klick auf den Home-Button/ Auge, um in das Portal zu gelangen)	35%	45%	10%	10%
Graphisch Browsen im Portal (Netzdarstellung)	30%	35%	20%	15%
Suche im Portal nach Webseiten und Tags	50%	30%	5%	15%
Möglichkeit in den Wissensnetzen (Ressourcen, Tags) anderer Benutzer zu browsen	25%	50%	10%	15%
Erweiterungen zu definieren, um mehr Informationen abspeichern zu können (eigene Formularfelder beim Speichern ähnlich dem BibTex-Feld)	20%	25%	30%	25%
Das Exportieren des Wissensnetzes als BibTex / HTML	45%	25%	15%	15%

Tabelle 9.10: Wichtigkeit der Funktionen in ELWMS.KOM

In Tabelle 9.11 ist die Gesamt-Einschätzung der Teilnehmer angegeben. Die Teilnehmer stimmten den Aussagen eher bis überwiegend zu, d.h. der Prototyp wurde insgesamt positiv angenommen, wenn

auch nicht von allen Teilnehmern. Die prozentuale Verteilung der Antworten ist ebenfalls in der Tabelle angegeben.

Sie haben ELWMS genutzt, um Recherchen durchzuführen. Bitte geben Sie an, inwieweit Sie den folgenden Aussagen zustimmen.	Stimme voll und ganz zu (6)	Stimme überwiegend zu (5)	Stimme eher zu (4)	Stimme eher nicht zu (3)	Stimme überwiegend nicht zu (2)	Stimme überhaupt nicht zu (1)	Mittelwert	Standardabweichung
Ich habe ELWMS für meine Recherchen gern genutzt.	30%	25%	30%	5%	5%	5%	4,55	1,39
ELWMS war sehr gut auf die Anforderungen der Recherche zugeschnitten.	30%	20%	15%	15%	15%	5%	4,20	1,64
ELWMS hat mir geholfen, bei der Recherche strukturierter vorzugehen als normalerweise.	35%	30%	10%	10%	10%	5%	4,55	1,57
Ich würde ELWMS meinen Freunden empfehlen, wenn diese Recherchen durchführen und ihre Webseiten verwalten müssen.	35%	30%	10%	10%	10%	5%	4,55	1,57
ELWMS war mir bei der Organisation der gefundenen Quellen (Webseiten) nützlich.	35%	30%	15%	5%	15%	0%	4,65	1,42
Die Qualität meiner Recherchen schätze ich durch die Nutzung von ELWMS besser ein.	25%	30%	5%	20%	15%	5%	4,15	1,63
Ich würde ELWMS gern für weitere Recherchen nutzen.	60%	5%	10%	10%	10%	5%	4,80	1,70

Tabelle 9.11: Meinung der Teilnehmer über ELWMS.KOM nach Ende der Nutzung

In den folgenden vier Tabellen sind die Fragen entsprechend der Herausforderungen des Ressourcenbasierten Lernens angeordnet, um zu zeigen, welche Effekte die Software auf diese hat. Auf Grund der eher geringen Teilnehmerzahl und der eher hohen Standardabweichung stellen diese Ergebnisse nur Trends dar. Die prozentuale Verteilung aller Antworten und ihre Standardabweichung ist in Tabelle B.10 im Anhang zu finden. Die Mittelwerte der Antworten zeigen einen eher positiven Effekt auf die Herausforderungen beim Ressourcen-basierten Lernen.

Frage	Mittelwert	Aussage	Effekt
Die Möglichkeit Ziele anzulegen hat mich bei der Planung meiner Recherchen unterstützt.	4,50	Stimme überwiegend zu	+
Wenn eine Webseite bereits von anderen Nutzern gespeichert wurde, hat mich dies bei der Einschätzung der Qualität und Vertrauenswürdigkeit der Webseite unterstützt.	4,46	Stimme eher zu	+
Die Planung der Recherche hat mir während der Durchführung bei der Beurteilung der Relevanz geholfen.	4,00	Stimme eher zu	+
Die Vorschläge in der Sidebar und im Portal haben mir geholfen, weitere relevante Informationen zu finden.	3,44	Stimme eher nicht zu	-
Die Ziele in der Sidebar haben mir geholfen, eine begonnene Recherche weiter zu planen und wenn nötig meine Suchstrategie anzupassen.	4,44	Stimme eher zu	+
Die Ziele und deren zugeordnete Webseiten haben mich dabei unterstützt, mich an eine frühere oder unterbrochene Recherche zu erinnern und schnelle Anknüpfungspunkte zu haben.	4,53	Stimme überwiegend zu	+

Tabelle 9.12: Fragen bzgl. struktureller Desorientierung

In Tabelle 9.12 sind die Fragen bezüglich der strukturellen Desorientierung gruppiert. Ein + in der letzten Spalte kennzeichnet einen positiven Effekt. Die Antworten zeigen, dass der Prototyp die Teilnehmer unterstützen kann, bei der Recherche strukturierter vorzugehen. Vor allem die Zielmanagementfunktio-

nen helfen bei der Reflektion des Vorgehens und des Fortschritts. Die Frage nach den vorgeschlagenen Ressourcen beim Rechercheprozess wurde leicht negativ beantwortet. Allerdings konnten den Teilnehmern in der Evaluation auch nur wenige Ressourcen vorgeschlagen werden, da die Teilnehmer jeweils sehr spezielle Themen recherchiert haben und es damit kaum Überschneidungen zwischen den Wissensnetzen verschiedener Nutzer gab und daher auch wenige weitere relevante Ressourcen.

In Tabelle 9.13 sind die Fragen bzgl. der konzeptuellen Desorientierung angeordnet. Das typisierte Tagging unterstützt die Nutzer des Prototypen dabei, die neuen Ressourcen in ihre bereits vorhandene Wissensnetzstruktur einzubauen, d.h. die Nutzer werden dazu angeregt, über die Zusammenhänge zwischen der neuen Ressource und ihrem vorhandenen Wissensnetz nachzudenken. Dabei bauen sie aktiv ein Wissensnetz auf. Beim Browsen im Wissensnetz haben die Teilnehmer teilweise neue Zusammenhänge entdecken können.

Frage	Mittelwert	Aussage	Effekt
Beim Taggen habe ich mir überlegt zu welchem Thema ich die Webseite einsortieren kann.	4,80	Stimme überwiegend zu	+
Die typisierten Tags gaben mir die Möglichkeit, Zusammenhänge zwischen einer gefundenen Webseite und der Recherche gut im Wissensnetz abzuspeichern.	4,58	Stimme überwiegend zu	+
Tags, die beim Tagging vorgeschlagen wurden, haben das Tagging gut unterstützt.	4,47	Stimme eher zu	+
Beim Browsen im Wissensnetz habe ich neue Zusammenhänge entdeckt.	3,85	Stimme eher zu	+

Tabelle 9.13: Fragen bzgl. konzeptueller Desorientierung

In Tabelle 9.14 sind die Fragen bzgl. der kognitiven Mehrbelastung gruppiert. Die Antworten zeigen einen positiven Effekt des Prototypen. Die Einbettung des Werkzeugs für das Ressourcenmanagement in den Webbrowser reduziert die Anzahl der Wechsel zwischen dem Webbrowser und anderen Anwendungen. Der Aufwand zum Speichern der Ressourcen im Wissensnetz wird eher gering eingeschätzt bzw. hat einen angemessenen Aufwand, so dass sich die Nutzer eher auf das tatsächliche Lesen der Webseiten konzentrieren können. Zusätzlich reduziert das Einblenden der Software im Webbrowser die Gefahr, sich von nicht relevanten Webseiten ablenken zu lassen.

Frage	Mittelwert	Aussage	Effekt
Durch die Möglichkeit, gefundene Informationen direkt mit der Sidebar abzuspeichern, konnte ich Wechsel zwischen Browser und anderen Anwendungen reduzieren bzw. vermeiden.	5,05	Stimme überwiegend zu	+
Das Umordnen im Wissensnetz und Einfügen neuer Informationen ist einfach.	4,13	Stimme eher zu	+
Die eingeblendete Sidebar hat die Gefahr reduziert, dass ich mich von anderen Webseiten ablenken lasse, die für die aktuelle Recherche nicht relevant sind.	4,00	Stimme eher zu	+
Während einer Recherche konzentriere ich mich auf das Lesen der gefundenen Informationen.	4,30	Stimme eher zu	+
Wenn ich an das Wissensnetz denke, das während meiner Recherche entstanden ist, finde ich den Aufwand des Abspeicherns der Webseiten angemessen.	4,81	Stimme überwiegend zu	+

Tabelle 9.14: Fragen bzgl. kognitiver Mehrbelastung

In Tabelle 9.15 sind die Fragen hinsichtlich des allgemeinen Ressourcenmanagements angeordnet. Die Teilnehmer gaben an, dass sie die zu einem früheren Zeitpunkt besuchte Webseiten überwiegend wie-

derfinden konnten. Allerdings wurde das entstandene Wissensnetz als eher unübersichtlich eingeschätzt. Hier gibt es durch die Integration der Pflegefunktionen Optimierungspotential. In einer anderen Frage gaben einige der Teilnehmer bereits an, dass sie überflüssige Tags eingegeben haben. Die Reduzierung solcher überflüssigen Tags kann die Übersichtlichkeit erhöhen. Die Teilnehmer gingen beim Tagging eher unstrukturiert vor, so dass Pflege- und Umordnungsfunktionen auch dabei helfen können, die Struktur im Wissensnetz zu erhöhen.

Die Antworten zu den Fragen, die vor allem im Zusammenhang mit den Wissensnetzen anderer Teilnehmer stehen, zeigen, dass es einen hohen Bedarf für die im Konzept-Kapitel beschriebenen Community-Funktionen gibt.

Frage	Mittelwert	Aussage	Effekt
Das Taggen mit typisierten Tags hat mir geholfen, Seiten wiederzufinden, die ich zu einem früheren Zeitpunkt bereits besucht habe.	4,53	Stimme überwiegend zu	+
Ich habe im Wissensnetz gefunden, wonach ich gesucht habe.	3,00	Stimme eher nicht zu	-
Ich finde das Wissensnetz, das bei mir entstanden ist, unübersichtlich.	3,80	Stimme eher zu	-
Ich konnte von den Webseiten und Wissensnetzen der anderen Nutzer profitieren.	3,06	Stimme eher nicht zu	-
Der Austausch von Recherche-Ergebnissen mit anderen Nutzern von ELWMS war umständlich.	3,08	Stimme eher nicht zu	+
Das Zurechtfinden in Wissensnetzen anderer Nutzer war schwierig.	3,80	Stimme eher zu	-

Tabelle 9.15: Fragen zum Aufwand des Ressourcenmanagements

9.4 Zusammenfassung

In diesem Kapitel wurden die im Rahmen dieser Arbeit durchgeführten Evaluationen mit dem Prototypen ELWMS.KOM bzw. dessen Vorläufern beschrieben. Die Ergebnisse der Evaluationen zeigen durchweg eine positive Annahme des Werkzeugs beim Ressourcen-basierten Lernen, wenn auch nicht jeder Nutzer der Software komplett überzeugt werden konnte. In der im letzten Abschnitt beschriebenen Studie konnte gezeigt werden, dass die Software das Potential hat, die Herausforderungen des Ressourcen-basierten Lernens zu reduzieren. Die Ergebnisse zeigen aber auch, an welchen Stellen der Prototyp in zukünftigen Forschungsarbeiten z.B. durch Integration der Community-Funktionen noch weiter optimiert werden kann.

10 Zusammenfassung und Ausblick

Dieses Kapitel schließt die Arbeit mit einer Zusammenfassung der Hauptbeiträge dieser Arbeit und einem Ausblick für weitere Arbeiten, deren Ansatzpunkte sich im Rahmen dieser Arbeit ergeben haben, ab.

10.1 Zusammenfassung

Ziel dieser Arbeit war die Konzeption, Entwicklung und Evaluation eines Werkzeugs zur Unterstützung von Lernenden beim Ressourcen-basierten Lernen. Insbesondere sollte das Werkzeug eine geeignete Möglichkeit zu Verwaltung der beim Lernen verwendeten webbasierten Ressourcen bieten, um die Herausforderungen des Ressourcen-basierten Lernens zu adressieren. Dieses Ziel wurde mittels folgender Beiträge erreicht:

Es wurde ein Modell für das Ressourcen-basierte Lernen entwickelt, bestehend aus den Prozessschritten Planen & Reflektieren, Suchen, Annotieren & Organisieren, Nutzen und Weitergeben. Die Entwicklung des Modells geschah auf Basis einer Analyse des Ressourcen-basierten Lernens in der Literatur und einer im Rahmen dieser Arbeit durchgeführten Studie.

Es wurden technische und funktionale Anforderungen an ein Werkzeug bestimmt, um alle im Modell definierten Prozessschritte des Ressourcen-basierten Lernens zu unterstützen. Ausgehend von diesen Anforderungen wurde ein Konzept für das Werkzeug entworfen.

Das Konzept sieht ein Wissensnetz als Basis vor, das mittels semantischem Tagging durch die Lernenden erstellt und erweitert wird. Tagging stellt eine breit akzeptierte Möglichkeit zur Verwaltung von beliebigen Ressourcen im Web dar, hat allerdings nur eingeschränkte Ausdrucksmöglichkeiten. Semantisches Tagging erweitert diese Ausdrucksmöglichkeiten um Elemente der formalen Wissensorganisation wie Begriffsordnungen.

Das typisierte Tagging als eine Form des semantischen Taggings wurde im Rahmen dieser Arbeit eingeführt. Mittels einer Analyse der Literatur und verschiedener im Rahmen dieser Arbeit durchgeführten Studien wurden sechs verschiedene Tag-Typen als wichtig für das Szenario des Ressourcen-basierten Lernens identifiziert. In den in dieser Arbeit durchgeführten Nutzerstudien zeigte sich, dass das Konzept des typisierten Taggings von den Benutzern angenommen wird und dass die definierten Tag-Typen sinnvoll nutzbar sind.

Zur automatischen Erkennung des Typs eines vom Benutzer eingegebenen Tags wurde in dieser Arbeit ein Algorithmus konzipiert, implementiert und auf verschiedenen Korpora evaluiert. Der Algorithmus klassifiziert die Tags ähnlich präzise wie verwandte Algorithmen, kann aber im Gegensatz zu diesen, die nur englischsprachige Tags klassifizieren, auch deutschsprachige Tags verarbeiten. Aufgrund des guten Laufzeitverhaltens des Algorithmus kann bereits während der Eingabe eines Tags durch den Nutzer der Tag-Typ identifiziert werden.

Ausgehend vom Konzept für das Unterstützungswerkzeug wurde die Plattform ELWMS.KOM entwickelt. Da die Verwaltung der Ressourcen direkt im Webbrowser stattfinden soll, wurde als Kernkomponente von ELWMS.KOM ein Add-on für den Webbrowser Mozilla Firefox entwickelt. Des Weiteren besteht die Plattform aus einem Backend zur Verwaltung des Wissensnetzes und einem webbasierten Frontend.

Schließlich wurde das Werkzeug in mehreren Benutzerstudien hinsichtlich seines Einsatzes im Ressourcen-basierten Lernen evaluiert. Diese Studien konnten zeigen, dass das Werkzeug die Herausforderungen des Ressourcen-basierten Lernens reduzieren kann und damit gut geeignet ist, das Ressourcen-basierte Lernen zu unterstützen.

10.2 Ausblick

Für weitere Forschungsarbeiten gibt es verschiedene Anknüpfungspunkte in ELWMS.KOM: In den Studien konnte festgestellt werden, dass das von den Lernenden erstellte Wissensnetz nach einer gewissen Nutzungszeit teilweise unübersichtlich wird. Zwar helfen die Tag-Typen beim typisierten Tagging den Lernenden dabei, mehr Struktur in das Wissensnetz zu bringen als bei Tagging-Systemen, die keine explizite Semantik unterstützen, aber dennoch werden weitere Mechanismen benötigt, die das Wissensnetz noch besser strukturieren. Beispielsweise gaben die Teilnehmer an den Studien zur Beobachtung der Langzeitnutzung an, dass sie häufiger Tags vergeben, die später nicht mehr gebraucht werden. Daher sollten in weiteren Forschungsarbeiten Mechanismen untersucht werden, die die Lernenden beim "Aufräumen" ihres Wissensnetzes unterstützen.

Des Weiteren gibt es Ansatzpunkte für die Optimierung von ELWMS.KOM, beispielsweise die Implementierung von Add-ons für weitere Webbrowser. In dieser Arbeit wurden hauptsächlich textbasierte Webressourcen adressiert, doch im Allgemeinen werden auch PDF-, Word-Dokumente oder andere Offline-Dokumente von den Lernenden verwendet, so dass Plug-ins für diese Anwendungen den manuellen Aufwand beim Verwalten der Ressourcen weiter senken können.

Eine weitere Herausforderung an ELWMS.KOM stellen die Community-Anforderungen dar. Denn Lernen findet oftmals in Lerngruppen statt, die gemeinsame Lernziele bearbeiten und Ressourcen recherchieren wollen. Im Konzept dieser Arbeit wurden bereits Ansätze für Community-Funktionalitäten erarbeitet, aber noch nicht implementiert und in die Plattform integriert. Diese Community-Funktionen werden im laufenden Projekt "CROKODIL" weiterentwickelt, implementiert und evaluiert.

Literaturverzeichnis

- [1] DIN 1463-1: Erstellung und Weiterentwicklung von Thesauri; Einsprachige Thesauri, 1987.
- [2] Fabian Abel, Ricardo Kawase, Daniel Krause, Patrick Siehndel, and Nicole Ullmann. The Art of Multi-faceted Tagging – Interweaving spatial annotations, categories, meaningful URIs and tags. In *6th Int. Conf. on Web Information Systems and Technologies (WEBIST)*, Valencia, Spain, April 2010.
- [3] Advanced Distributed Learning Initiative ADL. SCORM 2004 4th Edition Version 1.1 Overview, 2004. URL <http://www.adlnet.gov/Technologies/scorm/SCORMSDocuments/20044thEdition/>. Zuletzt abgerufen am 09.11.2010.
- [4] Hend S. Al-Khalifa. *Automatic document-level semantic metadata annotation using folksonomies and domain ontologies*. PhD thesis, University of Southampton, 2007.
- [5] American Library Association. Presidential Committee on Information Literacy: Final Report. Technical report, ALA, Chicago, IL, 1989.
- [6] Morgan Ames and Mor Naaman. Why We Tag: Motivations for Annotation in Mobile and Online Media. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '07, pages 971–980, New York, NY, USA, 2007. ACM.
- [7] Graham Attwell. Personal Learning Environments - the future of eLearning? *eLearning Papers*, 2 (1):1–8, 2007.
- [8] Susan Avery, Randy Brooks, James Brown, Paul Dorsey, and Michael O’Conner. Personal knowledge management: framework for integration and partnerships. In *Proceedings of ASCUE Conference*, 2001.
- [9] Andrea Back, Horst Baumgartner, Norbert Gronau, and Klaus Tochtermann, editors. *Web 2.0 in der Unternehmenspraxis. Grundlagen, Fallstudien und Trends zum Einsatz von Social Software*. Oldenburg Wissenschaftsverlag, 2008.
- [10] Walter R. Baets. *Organization Learning and Knowledge Technologies in a Dynamic Environment*. Kluwer Academic Publishers Norwell, MA, USA, 1998.
- [11] Ricardo Baeza-Yates and Berthier Ribeiro-Neto. *Modern Information Retrieval*. Addison Wesley, 1999.
- [12] Matthias Bauer, Ronald Maier, and Stefan Thalmann. Metadata Generation for Learning Objects: An Experimental Comparison of Automatic and Collaborative Solutions. In Michael H. Breitner, Franz Lehner, Jörg Staff, and Udo Winand, editors, *E-Learning 2010*, pages 181–195. Physica-Verlag HD, 2010.

-
- [13] Jöran Beel and Bela Gipp. Enhancing Search Applications by Utilizing Mind Maps. In *Proceedings of the 21st ACM Conference on Hypertext and Hypermedia*, HT '10, pages 303–304, New York, NY, USA, 2010. ACM.
- [14] Heiko Beier. Vom Wort zum Wissen. Semantische Netze als Mittel gegen die Informationsflut. *Information Wissenschaft und Praxis*, 55(3):133–138, 2004.
- [15] Bastian F. Benz. *Improving the Quality of E-Learning by Enhancing Self-Regulated Learning. A Synthesis of Research on Self-Regulated Learning and an Implementation of a Scaffolding Concept*. PhD thesis, Technische Universität Darmstadt, 2010.
- [16] Kerstin Bischoff, Claudiu S. Firan, Cristina Kadar, and Wolfgang Nejdl und Raluca Paiu. Automatically Identifying Tag Types. In Ronghuai Huang, Qiang Yang, Jian Pei, João Gama, Xiaofeng Meng, and Xue Li, editors, *Advanced Data Mining and Applications*, volume 5678 of *Lecture Notes in Computer Science*, pages 31–42. Springer Berlin / Heidelberg, 2009.
- [17] Christian Bizer, Jens Lehmann, Georgi Kobilarov, Sören Auer, Christian Becker, Richard Cyganiak, and Sebastian Hellmann. DBpedia – A Crystallization Point for the Web of Data. *Journal of Web Semantics: Science, Services and Agents on the World Wide Web*, Issue 7:154–165, 2009.
- [18] Simone Braun, Andreas Schmidt, Andreas Walter, and Valentin Zacharias. Von Tags zu semantischen Beziehungen: kollaborative Ontologiereifung. In Birgit Gaiser, Thorsten Hampel, and Stefanie Panke, editors, *Good Tags - Bad Tags: Social Tagging in der Wissensorganisation*. Waxmann Münster, 2008.
- [19] Patricia Senn Breivik. Information literacy. *Bulletin of the Medical Library Association*, 79(2): 226–229, 1991.
- [20] Robert L. Brennan and Dale J. Prediger. Coefficient Kappa: Some uses, misuses, and alternatives. *Educational and Psychological Measurement*, 41:687–699, 1981.
- [21] Sergey Brin and Lawrence Page. The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems*, 30(1-7):107 – 117, 1998.
- [22] John Brooke. SUS: A Quick and Dirty Usability Scale. In Patrick W. Jordan, Bruce Thomas, Bernard A. Weerdmeester, and Ian Lyall McClelland, editors, *Usability Evaluation in Industry*. London: Taylor & Francis, 1996.
- [23] Christopher Brooks, Scott Bateman, Wengang Liu, Gordon McCalla, Jim Greer, Dragan Gasevic, Timmy Eap, Griff Richards, Khaled Hammouda, Shady Shehata, Mohamed Kamel, Fakhri Karray, and Jelena Jovanović. Issues and directions with educational metadata. In *Third Annual International Scientific Conference of the Learning Object Repository Research Network*, Montreal, 2006.
- [24] Tony Buzan and Barry Buzan. *The mind map book*. BBC Active, 3 edition, 2006.
- [25] Alberto J. Cañas and Joseph D. Novak. Re-examining the foundations for effective use of concept maps. In *Concept Maps: Theory, Methodology, Technology. Proceedings of the Second International Conference on Concept Mapping*, volume 1, pages 494–502, 2006.

-
- [26] Kris Cardinaels, Michael Meire, and Erik Duval. Automating metadata generation: the simple indexing interface. In *Proceedings of the 14th international conference on World Wide Web, WWW '05*, pages 548–556, New York, NY, USA, 2005. ACM.
- [27] Vinesh Chandra and Christina Chalmers. Blogs, wikis and podcasts: Collaborative knowledge building tools in a Design and Technology course. *Journal of Learning Design*, 3(2):35–49, 2010.
- [28] Wen-hao Chen, Yi Cai, Ho-fung Leung, and Qing Li. Generating ontologies with basic level concepts from folksonomies. *Procedia Computer Science*, 1(1):573 – 581, 2010.
- [29] Ricky K.F. Cheong and Eric Tsui. The roles and values of personal knowledge management: An exploratory study. *VINE*, 40(2):204–227, 2010.
- [30] Namyoun Choi, Il-Yeol Song, and Hyoil Han. A survey on ontology mapping. *ACM SIGMOD Record*, 35(3):34–41, September 2006.
- [31] Stijn Christiaens. Metadata Mechanisms: From Ontology to Folksonomy ... and Back. In Robert Meersman, Zahir Tari, and Pilar Herrero, editors, *On the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops*, volume 4277 of *Lecture Notes in Computer Science*, pages 199–207. Springer Berlin / Heidelberg, 2006.
- [32] Timothy J. Cleary and Barry J. Zimmerman. Self-regulation empowerment program: A school-based program to enhance self-regulated and self-motivated cycles of student learning. *Psychology in the Schools*, 41(5):537–550, 2004.
- [33] Jacob Cohen. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, 20(1):37, 1960.
- [34] Jeff Conklin. Hypertext: A survey and introduction. *IEEE Computer*, 20(9):17–41, 1987.
- [35] Anthony V. D’Antoni, Genevieve Pinto Zipp, Valerie G. Olson, and Terrence F. Cahill. Does the mind map learning strategy facilitate information retrieval and critical thinking in medical students? *BMC Medical Education*, 10(1):61–72, 2010.
- [36] Thomas H. Davenport and Laurence Prusak. *Wenn Ihr Unternehmen wüßte, was es alles weiß ... das Praxishandbuch zum Wissensmanagement (Aus dem Amerik. übers.)*. Verl. Moderne Industrie, Landsberg/Lech, 1998.
- [37] Stefan Decker and Martin Frank. The social semantic desktop. In *Proceedings of the WWW2004 Workshop Application Design, Development and Implementation Issues in the Semantic Web*, 2004.
- [38] Wissenschaftlicher Rat der Dudenredaktion, editor. *DUDEN Fremdwörterbuch*. Duden Verlag, Mannheim, Leipzig, Zürich, Wien, 5 edition, 1990.
- [39] Ullrich Dittler, Michael Kindt, and Christine Schwarz, editors. *Online-Communities als soziale Systeme. Wikis, Weblogs und Social Software im E-Learning*. Waxmann Münster, 2007.
- [40] Michael Douma and Greg Ligierko. Creating Online Mind Maps and Concept Maps. In *25th Annual Conference on Distance Teaching & Learning*, pages 1–8, 2009.

-
- [41] Sarah Edwards and Nick Cooper. Mind mapping as a teaching resource. *The Clinical Teacher*, 7(4):236–239, 2010.
- [42] Andreas Faatz. *Ein Verfahren zur Anreicherung fachgebietsspezifischer Ontologien durch Begriffsvorschläge*. PhD thesis, TU Darmstadt, 2004.
- [43] Giorgos Flouris, Dimitris Manakanatas, Haridimos Kondylakis, Dimitris Plexousakis, and Grigoris Antoniou. Ontology change: classification and survey. *The Knowledge Engineering Review*, 23(02):117–152, 2008.
- [44] Andrés Garca-Silva, Martin Szomszor, Harith Alani, and Oscar Corcho. Preliminary results in tag disambiguation using dbpedia. In *Knowledge Capture (K-Cap’09)-Workshop on Collective Knowledge Capturing and Representation-CKCaR*, 2009.
- [45] Barbara Geyer-Hayden. Wissensmodellierung im Semantic Web. In Andreas Blumauer and Tassilo Pellegrini, editors, *Social Semantic Web*, X.media.press, pages 127–146. Springer Berlin Heidelberg, 2009.
- [46] Tony Gill, Anne J. Gilliland, Maureen Whalen, and Mary Woodley. *Introduction to Metadata*. Getty Research Insitute, Los Angeles, online edition, version 3.0 edition, 2008.
- [47] Scott A. Golder and Bernardo A. Huberman. Usage patterns of collaborative tagging systems. *Journal of Information Science*, 32(2):198–208, 2006.
- [48] Jane S. Grassley and Robin Bartoletti. Wikis and Blogs: Tools for Online Interaction. *Nurse Educator*, 34(5):209–213, 2009.
- [49] Joachim Griesbaum. *Mehrwerte des Kollaborativen Wissensmanagements in der Hochschullehre - Integration asynchroner netzwerkbasierter Szenarien des CSCL in der Ausbildung der Informationswissenschaft im Rahmen des K3-Projekts*. PhD thesis, Universität Konstanz, 2006.
- [50] Thomas R. Gruber. A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2):199–220, 1993.
- [51] Janis Grundspenkis. Agent based approach for organization and personal knowledge modelling: knowledge management perspective. *Journal of Intelligent Manufacturing*, 18:451–457, 2007.
- [52] Martin Gudgin, Marc Hadley, Noah Mendelsohn, Jean-Jacques Moreau, Henrik Frystyk Nielsen, Anish Karmarkar, and Yves Lafon (W3C XML Protocol Working Group). SOAP - W3C Recommendation 27 April 2009, 2007. URL <http://www.w3.org/TR/soap12/>. Zuletzt abgerufen am 16.03.2011.
- [53] Georg Güntner, Rolf Sint, and Rupert Westenthaler. Ein Ansatz zur Unterstützung traditioneller Klassifikation durch Social Tagging. In Birgit Gaiser, Thorsten Hampel, and Stefanie Panke, editors, *Good Tags - Bad Tags: Social Tagging in der Wissensorganisation*. Waxmann Münster, 2008.
- [54] Hartmut Häcker and Kurt H. Stapf, editors. *Dorsch: Psychologisches Wörterbuch (13., überarbeitete und erweiterte Auflage)*. Hans Huber, Bern, 1998.

-
- [55] Heiko Haller and Andreas Abecker. Designing a Knowledge Mapping Tool for Knowledge Workers. In Rossitza Setchi, Ivan Jordanov, Robert Howlett, and Lakhmi Jain, editors, *Knowledge-Based and Intelligent Information and Engineering Systems*, volume 6276 of *Lecture Notes in Computer Science*, pages 660–669. Springer Berlin / Heidelberg, 2010.
- [56] Tony Hammond, Timo Hannay, Ben Lund, and Joanna Scott. Social bookmarking tools (I): A General Review. *D-Lib Magazine*, 11(4), 2005.
- [57] Michael Hannafin and Janette Hill. Resource-based learning. In J. Michael Spector, M. David Merrill, Jeroen van Merriënboer, and Marcy P. Driscoll, editors, *Handbook of research on educational communications and technology*, pages 525–536. Routledge/Taylor & Francis Group, 3 edition, 2007.
- [58] Markus Heckner, Michael Heilemann, and Christian Wolff. Personal Information Management vs. Resource Sharing: Towards a Model of Information Behaviour in Social Tagging Systems. In *Third International AAAI Conference on Weblogs and Social Media (ICWSM)*, pages 42–49, 2009.
- [59] Sandra Higgison. Your say: Personal knowledge management. *Knowledge Management Magazine*, 7(7):11–12, 2004.
- [60] Janette Hill and Michael Hannafin. Teaching and learning in digital environments: The resurgence of resource-based learning. *Educational Technology Research and Development*, 49:37–52, 2001.
- [61] Gail Hodge. Systems of Knowledge Organization for Digital Libraries: Beyond Traditional Authority Files. Technical report, The Digital Library Federation Council on Library and Information Resources, 2000.
- [62] Stefan Hoermann, Christoph Rensing, and Ralf Steinmetz. Wiederverwendung von Lernressourcen mittels Authoring by Aggregation im ResourceCenter. In Jörg M. Haake, Ulrike Lucke, and Djamshid Tavangarian, editors, *DeLFI 2005: 3. Deutsche e-Learning Fachtagung Informatik*, pages 153–164. GI-Edition, Sep 2005.
- [63] Andreas Hotho. Data Mining on Folksonomies. In Giuliano Armano, Marco de Gemmis, Giovanni Semeraro, and Eloisa Vargiu, editors, *Intelligent Information Access*, volume 301 of *Studies in Computational Intelligence*, pages 57–82. Springer Berlin / Heidelberg, 2010.
- [64] Andreas Hotho, Robert Jäschke, Dominik Benz, Miranda Grahland Beate Krause, Christoph Schmitz, and Gerd Stumme. Social Bookmarking am Beispiel BibSonomy. In Andreas Blumauer and Tassilo Pellegrini, editors, *Social Semantic Web*, X.media.press, pages 363–391. Springer Berlin Heidelberg, 2009.
- [65] Willem-Olaf Huijsen, Samuël J. Driessen, and Jan W.M. Jacobs. Explicit conceptualizations for Knowledge Mapping. In *International Conference on Enterprise Information (ICEIS 2004)*, volume 3, pages 231–236, Porto, April 2004, 2004.
- [66] Institute of Electrical and Electronics Engineers. IEEE Standard for Learning Object Metadata 1484.12.1., 2002. URL http://ltsc.ieee.org/wg12/files/LOM_1484_12_1_v1_Final_Draft.pdf. Zuletzt abgerufen am 03.11.2010.

-
- [67] International Organization for Standardization. ISO 2788:1986 Documentation – Guidelines for the establishment and development of monolingual thesauri, 1986.
- [68] International Organization for Standardization. ISO/IEC 13250 Information Technology – Topic Maps, 200. URL http://www1.y12.doe.gov/capabilities/sgml/sc34/document/0322_files/iso13250-2nd-ed-v2.pdf. Zuletzt abgerufen am 18.01.2011.
- [69] International Organization for Standardization. ISO 15836:2009 Information and documentation – The Dublin Core metadata element set, Februar 2009. URL http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=52142. Zuletzt abgerufen am 15.11.2010.
- [70] Jon Iturrioz, Oscar Diaz, and Sergio F. Anzuola. Toward the Semantic Desktop: The seMouse Approach. *Intelligent Systems, IEEE*, 23(1):24–31, 2008.
- [71] Karl Ludwig Jüngst and Peter Strittmatter. Wissensstrukturdarstellung: Theoretische Ansätze und praktische Relevanz. *Unterrichtswissenschaft*, 23(3):194–207, 1995.
- [72] Christiane Kaufmann, editor. *WordNet - An Electronic Lexical Database*. Bradford Books, 1998.
- [73] Michael Kerres. Potenziale des Lernens im Internet: Fiktion oder Wirklichkeit. In Hilmar Hoffmann, editor, *Deutsch global? Neue Medien, eine Herausforderung für die deutsche Sprache?*, pages 170–195. DuMont, 2000.
- [74] Malte Kiesel. Kaukolu: Hub of the semantic corporate intranet. In *Proceedings of the Workshop SemWiki2006 - From Wiki to Semantics at the ESWC Conference*, 2006.
- [75] Margaret E. I. Kipp. Complementary or Discrete Contexts in Online Indexing: A Comparison of User, Creator, and Intermediary Keywords. *Canadian Journal of Information and Library Science*, 30(3), 2006.
- [76] Mathias Krebs, Christian Schmidt, Michael Henninger, Matthias Ludwig, and Wolfgang Müller. Are Wikis and Weblogs an Appropriate Approach to Foster Collaboration, Reflection and Students' Motivation? In Nicholas Reynolds and Márta Turcsányi-Szabó, editors, *Key Competencies in the Knowledge Society*, volume 324 of *IFIP Advances in Information and Communication Technology*, pages 200–209. Springer Boston, 2010.
- [77] Rainer Kuhlen. *Hypertext: Ein nicht-lineares Medium zwischen Buch und Wissensbank*. Springer Verlag Berlin, 1991.
- [78] Rainer Kuhlen. Information. In Rainer Kuhlen, Thomas Seeger, and Dietmar Strauch, editors, 5. Auflage des *Handbuchs Grundlagen von Information und Dokumentation*, pages 3–20. Saur-Verlag: München, 2004.
- [79] Jürg Kuster, Eugen Huber, Robert Lippmann, Alphons Schmid, Emil Schneider, Urs Witschi, and Roger Wüst. Methoden der Problemlösung. In *Handbuch Projektmanagement*, pages 338–375. Springer Berlin Heidelberg, 2008.

-
- [80] J. Richard Landis and Gary G. Koch. The measurement of observer agreement for categorical data. *Biometrics*, 33(1):159–174, 1977.
- [81] Kathy J. Lee. What Goes Around Comes Around: An analysis of del.icio.us as social space. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work, CSCW '06*, pages 191–194, New York, NY, USA, 2006. ACM.
- [82] Lasse Lehmann. *Lebenszyklusinformationen von Wissensdokumenten - Erfassung, Verwaltung und Validierung*. PhD thesis, TU Darmstadt, 2010.
- [83] Lasse Lehmann, Helge Fredrich, Christoph Rensing, Volker Zimmermann, and Ralf Steinmetz. Das Authoring Management System EXPLAIN zur ganzheitlichen Unterstützung des Erstellungsprozesses von Trainingsmedien und WBT. In Christian Eibl, Johannes Magenheimer, Sigrid Schubert, and Martin Wessner, editors, *DeLFI 2007: 5. e-Learning Fachtagung Informatik*, number P-111, pages 139–150, Köllen, Bonn, Sep 2007. GI, Lecture Notes in Informatics (LNI).
- [84] Allison Littlejohn, Anoush Margaryan, and Colin Milligan. Charting collective knowledge: Supporting self-regulated learning in the workplace. In *Ninth IEEE International Conference on Advanced Learning Technologies*, pages 208–212. IEEE, 2009.
- [85] Kaipeng Liu, Binxing Fang, and Weizhe Zhang. Ontology emergence from folksonomies. In *Proceedings of the 19th ACM international conference on Information and knowledge management, CIKM '10*, pages 1109–1118, New York, NY, USA, 2010. ACM.
- [86] Robert Lokaiczky. Clustering und Klassifikation von Personennamen (Diplomarbeit). Master's thesis, Universität Leipzig, 2006.
- [87] George Macgregor and Emma McCulloch. Collaborative Tagging as a Knowledge Organisation and Resource Discovery Tool. *Library Review*, 55(5):291–300, 2006.
- [88] John Makhoul, Francis Kubala, Richard Schwartz, and Ralph Weischedel. Performance measures for information extraction. In *Proceedings of DARPA Broadcast NewsWorkshop*, 1999.
- [89] Andreas Marchetti, Maurizio Tesconi, Francesco Ronzano, Marco Rosella, and Salvatore Minutoli. Semkey: A semantic collaborative tagging system. In *Workshop on Tagging and Metadata for Social Information Organization at WWW*, volume 7, pages 8–12, 2007.
- [90] Cameron Marlow, Mor Naaman, Danah Boyd, and Marc Davis. HT06, tagging paper, taxonomy, Flickr, academic article, to read. In *Proceedings of the Seventeenth Conference on Hypertext and Hypermedia*, pages 31–40. ACM Press, New York, 2006.
- [91] Norbert Meder. *Web-Didaktik: Eine neue Didaktik webbasierten, vernetzten Lernens*. Bertelsmann, 2006.
- [92] Peter Meier. *Visualisierung von Kommunikationsstrukturen für kollaboratives Wissensmanagement in der Lehre*. PhD thesis, Universität Konstanz, 2006. URL <http://kops.ub.uni-konstanz.de/volltexte/2006/1925/>. Zuletzt abgerufen am 03.11.2010.

-
- [93] Marek Meyer, Birgit Zimmermann, Christoph Rensing, and Ralf Steinmetz. An Interactive Tool for Supporting Modularization of SCORM-Based Learning Resources. In *Proceedings of ED-MEDIA 2007*, pages 3164–3171. AACE, Jun 2007.
- [94] Peter Mika. Ontologies are us: A unified model of social networks and semantics. *Web Semantics: Science, Services and Agents on the World Wide Web*, 5(1):5–15, 2007.
- [95] Matthias Müller-Prove. Modell und Anwendungsperspektive des Social Tagging. In Birgit Gaiser, Thorsten Hampel, and Stefanie Panke, editors, *Good Tags - Bad Tags: Social Tagging in der Wissensorganisation*, volume 47 of *Medien in der Wissenschaft*, pages 15–22. Waxmann Münster, 2008.
- [96] Anja Naumann. *Wissenserwerb und Informationssuche mit Hypertexten: Die Bedeutung von Strukturierung, Navigationshilfen und Arbeitsgedächtnisbelastung*. PhD thesis, TU Chemnitz, 2003.
- [97] John C. Nesbit and Olusola O. Adesope. Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research*, 76(3):413–448, 2006.
- [98] Anja Neumann, Wolfgang Gräber, and Sigmar-Olaf Tergan. ParIS - Visualizing Ideas and Information in a Resource-Based Learning Scenario. In Sigmar-Olaf Tergan and Tanja Keller, editors, *Knowledge and Information Visualization*, Lecture Notes in Computer Science, pages 256–281. Springer-Verlag, Berlin Heidelberg, 2005.
- [99] Ikujiro Nonaka and Hirotaka Takeuchi. *Die Organisation des Wissens - Wie japanische Unternehmen eine brachliegende Ressource nutzbar machen*. Campus-Verlag, Frankfurt am Main, 1997. Aus dem Engl. von Friedrich Mader.
- [100] Simon Overell, Börkur Sigurbjörnsson, and Roelof van Zwol. Classifying tags using open content resources. In *Proceedings of the Second ACM International Conference on Web Search and Data Mining*, WSDM '09, pages 64–73, New York, NY, USA, 2009. ACM.
- [101] Larry Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. Technical report, Stanford Digital Library Technologies Project, 1999.
- [102] Stefanie Panke and Birgit Gaiser. "With my head up in the clouds" - Social tagging aus Nutzersicht. In Birgit Gaiser, Thorsten Hampel, and Stefanie Panke, editors, *Good Tags - Bad Tags: Social Tagging in der Wissensorganisation*. Waxmann Münster, 2008.
- [103] Jiri Panyr. Thesauri, Semantische Netze, Frames, Topic Maps, Taxonomien, Ontologien - begriffliche Verwirrung oder konzeptionelle Vielfalt? In Ilse Harms, Heinz-Dirk Luckhardt, and Hans W. Giessen, editors, *Information und Sprache. Beiträge zu Informationswissenschaft, Computerlinguistik, Bibliothekswesen und verwandten Fächern. Festschrift für Harald H. Zimmermann*, pages 139–150. K. G. Saur München, 2006.
- [104] John C. Paolillo and Shashikant Penumarthy. The Social Structure of Tagging Internet Video on del.icio.us. In *40th Annual Hawaii International Conference on System Sciences, 2007. HICSS 2007*, pages 85–95. IEEE, 2007.

-
- [105] Alexandre Passant and Philippe Laublet. Meaning of a tag: A collaborative approach to bridge the gap between tagging and linked data. In *Proceedings of the WWW 2008 Workshop Linked Data on the Web (LDOW2008)*, Beijing, China, 2008.
- [106] Oren Patashnik. BIBTEXing, Februar 1988. URL <http://bibtexml.sourceforge.net/btxdoc.pdf>. Zuletzt abgerufen am 16.03.2011.
- [107] Isabella Peters. *Folksonomies: Indexing and retrieval in Web 2.0*. KG Saur Verlag GmbH & Co, 2009.
- [108] Gilbert Probst, Steffen Raub, and Kai Romhardt. *Wissen managen. Wie Unternehmen ihre wertvollste Ressource optimal nutzen*. Gabler, Wiesbaden, 5 edition, 2006.
- [109] Dennis Quan, David Huynh, and David Karger. Haystack: A Platform for Authoring End User Semantic Web Applications. In *The SemanticWeb - ISWC 2003*, volume 2870 of *Lecture Notes in Computer Science*, pages 738–753. Springer Berlin / Heidelberg, 2003.
- [110] M. Ross Quillian. *Semantic Information Processing*, chapter Semantic Memory, pages 227–270. MIT Press, 1968.
- [111] Emanuele Quintarelli, Andrea Resmini, and Luca Rosati. Facetag: Integrating bottom-up and top-down classification in a social tagging system. *Bulletin of the American Society for Information Science and Technology*, 33(5):10–15, 2007.
- [112] Glenda C. Rakes. Using the internet as a tool in a resource-based learning environment. *Educational Technology*, September-October:52–56, 1996.
- [113] Justus J. Randolph. Free-Marginal Multirater Kappa (multirater kappa free): An Alternative to Fleiss' Fixed-Marginal Multirater Kappa. In *Joensuu Learning and Instruction Symposium*, 2005.
- [114] Justus J. Randolph. Online Kappa Calculator. <http://justus.randolph.name/kappa>, 2008. Zuletzt abgerufen am 02.03.2011.
- [115] RDF Working Group, World Wide Web Consortium. Resource Description Framework (RDF), Feb 2004. URL <http://www.w3.org/standards/techs/rdf>. Zuletzt abgerufen am 17.01.2011.
- [116] Gabi Reinmann and Martin J. Eppler. *Wissenswege. Methoden für das persönliche Wissensmanagement*. Lernen mit Neuen Medien. Verlag Hans Huber, 2008.
- [117] Gabi Reinmann and Heinz Mandl. Wissensmanagement und Weiterbildung. In Rudolf Tippelt and Aiga Hippel, editors, *Handbuch Erwachsenenbildung/Weiterbildung*, pages 1049–1066. VS Verlag für Sozialwissenschaften, 4 edition, 2010.
- [118] Gabi Reinmann-Rothmeier. Wissen managen: Das Münchener Modell, 2001. URL <http://epub.ub.uni-muenchen.de/239/>. Zuletzt abgerufen am 03.11.2010.
- [119] Christoph Rensing, Sonja Bergsträsser, Tomas Hildebrandt, Marek Meyer, Birgit Zimmermann, Andreas Faatz, Lasse Lehmann, and Ralf Steinmetz. Re-Use and Re-Authoring of Learning Resources - Definitions and Examples. Technical Report KOM-TR-2005-02, TU Darmstadt - Multimedia Communications Lab, Darmstadt, Nov 2005. URL

http://www.kom.tu-darmstadt.de/fileadmin/Externer_Bereich/Publications/Technical-Reports/technical-report-kom-2005-02.pdf. Zuletzt abgerufen am 10.01.2011.

- [120] Jennifer Rowley. The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science*, 33(2):163, 2007.
- [121] RSS Advisory Board. RSS 2.0 Specification, 2009. URL <http://www.rssboard.org/rss-specification>. Zuletzt abgerufen am 26.01.2011.
- [122] Michael F. Ruffini. Using E-Maps to Organize and Navigate Online Content Using E-Maps to Organize and Navigate Online Content. *Interactive Educational Multimedia*, 16:87–98, 2008.
- [123] Steve Ryan, Bernard Scott, Howard Freeman, and Daxa Patel. *The virtual university: The Internet and resource-based learning*. London: Kogan Press, 2000.
- [124] Leo Sauermann. *The Gnowsiss Semantic Desktop approach to Personal Information Management*. PhD thesis, TU Kaiserslautern, 2009.
- [125] Andreas Schmidt. *Situationsbewusste Informationsdienste für das arbeitsbegleitende Lernen*. PhD thesis, Universität Karlsruhe, 2009.
- [126] Jan Schmidt. Potenziale von Social Software für Bildungsportale. In Birgit Gaiser, Friedrich W. Hesse, and Monika Lütke-Entrup, editors, *Bildungsportale–Potenziale und Perspektiven netzbasierter Bildungsressourcen*, pages 219–233. Oldenbourg Wissenschaftsverlag, 2007.
- [127] Bernhard Schmitz and Bettina S. Wiese. New perspectives for the evaluation of training sessions in self-regulated learning: Time-series analyses of diary data. *Contemporary Educational Psychology*, 31(1):64–96, 2006.
- [128] Philipp Scholl, Bastian Benz, Doreen Böhnstedt, Christoph Rensing, Ralf Steinmetz, and Bernhard Schmitz. Einsatz und Evaluation eines Zielmanagement-Werkzeugs bei der selbstregulierten Internet-Recherche. In Silke Seehusen, Ulrike Lucke, and Stefan Fischer, editors, *DeLFI 2008: 6. e-Learning Fachtagung Informatik*, number P-132, pages 125–136, Köllen, Bonn, 2008. GI, Lecture Notes in Informatics (LNI).
- [129] Philipp Scholl, Bastian F. Benz, Doreen Böhnstedt, Christoph Rensing, Bernhard Schmitz, and Ralf Steinmetz. Implementation and Evaluation of a Tool for Setting Goals in Self-Regulated Learning with Web Resources. In Ulrike Cress, Vania Dimitrova, and Marcus Specht, editors, *Learning in the Synergy of Multiple Disciplines, EC-TEL 2009*, volume LNCS Vol 5794, pages 521–534. Springer-Verlag Berlin Heidelberg, 2009.
- [130] Philipp Scholl, Renato Domínguez García, Doreen Böhnstedt, Christoph Rensing, and Ralf Steinmetz. Towards Language-Independent Web Genre Detection. In *WWW '09: Proceedings of the 18th international conference on World Wide Web*, pages 1157–1158, New York, NY, USA, 2009. ACM.

-
- [131] Philipp Scholl, Doreen Böhnstedt, Renato Domínguez García, Christoph Rensing, and Ralf Steinmetz. Extended Explicit Semantic Analysis for Calculating Semantic Relatedness of Web Resources. In Martin Wolpers, Paul A. Kirschner, Maren Scheffel, Stefanie Lindstädt, and Vania Dimitrova, editors, *Sustaining TEL: From Innovation to Learning and Practice Proceedings of EC-TEL 2010*, volume Lecture Notes in Computer Science 6383, pages 324–339. Springer Verlag, 2010.
- [132] Ulrik Schroeder. Web-Based Learning - Yes We Can! In Marc Spaniol, Qing Li, Ralf Klamma, and Rynson Lau, editors, *Advances in Web Based Learning - ICWL 2009*, volume 5686 of *Lecture Notes in Computer Science*, pages 25–33. Springer Berlin / Heidelberg, 2009.
- [133] Ulrik Schroeder and Christian Spannagel. Lernen mit Web-2.0-Anwendungen. *Navigationen. Zeitschrift für Medien- und Kulturwissenschaften*, 8(1):59–79, 2008.
- [134] Shilad Sen, Shyong K. Lam, Al Mamunur Rashid, Dan Cosley, Dan Frankowski, Jeremy Osterhouse, F. Maxwell Harper, and John Riedl. tagging, communities, vocabulary, evolution. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work, CSCW '06*, pages 181–190, New York, NY, USA, 2006. ACM.
- [135] Kuei-Ping Shih, Tai-Chien Kao, Chih-Yung Chang, and Hung-Chang Chen. A study of self-regulated learning in high school students' english learning with system support. In *First IEEE International Conference on Ubi-Media Computing*, pages 296–301. IEEE, 2008.
- [136] Pavel Shvaiko and Jérôme Euzénat. A Survey of Schema-Based Matching Approaches. In Stefano Spaccapietra, editor, *Journal on Data Semantics IV*, volume 3730 of *Lecture Notes in Computer Science*, pages 146–171. Springer Berlin / Heidelberg, 2005.
- [137] Melody Siadaty, Jelena Jovanović, Dragan Gašević, Zoran Jeremić, and Teresa Holocher-Ertl. Leveraging Semantic Technologies for Harmonization of Individual and Organizational Learning. In Martin Wolpers, Paul Kirschner, Maren Scheffel, Stefanie Lindstaedt, and Vania Dimitrova, editors, *Sustaining TEL: From Innovation to Learning and Practice*, volume 6383 of *Lecture Notes in Computer Science*, pages 340–356. Springer Berlin / Heidelberg, 2010.
- [138] Miguel Á. Sicilia, Elena García, Carmen Pagés, José J. Martínez, and José M. Gutiérrez. Complete Metadata Records in Learning Object Repositories: Some Evidence and Requirements. *International Journal of Learning Technology*, 1(4):411–424, 2005.
- [139] Börkur Sigurbjörnsson and Roelof van Zwol. Flickr tag recommendation based on collective knowledge. In *Proceeding of the 17th international conference on World Wide Web, WWW '08*, pages 327–336, New York, NY, USA, 2008. ACM.
- [140] Börkur Sigurbjörnsson and Roelof van Zwol. Tagexplorer: Faceted Browsing of Flickr Photos. Technical report, Yahoo! Research, 2010.
- [141] Rashmi Sinha. A cognitive analysis of tagging, 2005. URL http://www.rashmisinha.com/archives/05_09/tagging-cognitive.html. Zuletzt abgerufen am 03.01.2011.

-
- [142] John F. Sowa. *Encyclopedia of Artificial Intelligence*, volume 2, chapter Semantic networks, pages 1493–1511. John Wiley, New York, 1992.
- [143] John F. Sowa. *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Brooks/Cole Publishing Co., Pacific Grove, CA, 2000.
- [144] Achim Steinacker. *Medienbausteine für web-basierte Lernsysteme*. PhD thesis, Technische Universität Darmstadt, 2002.
- [145] Wolfgang G. Stock and Mechtild Stock. *Wissensrepräsentation. Informationen auswerten und bereitstellen*. Oldenbourg Verlag München, 2008.
- [146] Alexander Stocker and Klaus Tochtermann. *Wissenstransfer mit Wikis und Weblogs: Fallstudien zum erfolgreichen Einsatz von Web 2.0 in Unternehmen*. Gabler Verlag, 2010.
- [147] Rudi Studer, V. Richard Benjamins, and Dieter Fensel. Knowledge Engineering: Principles and methods. *Data & Knowledge Engineering*, 25:161–197, 1998.
- [148] Sigmar-Olaf Tergan. Hypertext und Hypermedia: Konzeption, Lernmöglichkeiten, Lernprobleme und Perspektiven. In Ludwig J. Issing and Paul Klimsa, editors, *Information und Lernen mit Multimedia und Internet: Lehrbuch für Studium und Praxishandbuch*, chapter Hypertext und Hypermedia: Konzeption, Lernmöglichkeiten, Lernprobleme und Perspektiven, pages 99–113. Verlagsguppe Beltz, 3 edition, 2002.
- [149] Sigmar-Olaf Tergan. Digital Concept Maps for Managing Knowledge and Information. In Sigmar-Olaf Tergan and Tanja Keller, editors, *Knowledge and Information Visualization*, volume 3426 of *Lecture Notes in Computer Science*, pages 185–204. Springer Berlin / Heidelberg, 2005.
- [150] Giovanni Tummarello, Christian Morbidoni, and Michele Nucci. Enabling Semantic Web Communities with DBin: An Overview. In Isabel Cruz, Stefan Decker, Dean Allemang, Chris Preist, Daniel Schwabe, Peter Mika, Mike Uschold, and Lora Aroyo, editors, *The Semantic Web - ISWC 2006*, volume 4273 of *Lecture Notes in Computer Science*, pages 943–950. Springer Berlin / Heidelberg, 2006.
- [151] Mark van Harmelen. Personal Learning Environments. In *Proceedings of the Sixth IEEE International Conference on Advanced Learning Technologies*, pages 815–816, 2006.
- [152] Thomas Vander Wal. Folksonomy Coinage and Definition, 2007. URL <http://vanderwal.net/folksonomy.html>. Zuletzt abgerufen am 13.01.2011.
- [153] Jakob Voß. Vom Social Tagging zum Semantic Tagging. In Birgit Gaiser, Thorsten Hampel, and Stefanie Panke, editors, *Good Tags - Bad Tags: Social Tagging in der Wissensorganisation*. Waxmann Münster, 2008.
- [154] Max Völkel and Andreas Abecker. Cost-Benefit Analysis for the Design of Personal Knowledge Management Systems. In José Cordeiro and Joaquim Filipe, editors, *ICEIS 2008 - Proceedings of the Tenth International Conference on Enterprise Information Systems*, volume AIDSS, pages 95–105, Barcelona, Spain, 2008.

-
- [155] Andreas Wagner and Marc Rössler. WALU - Eine Annotations- und Lern-Umgebung für semantisches Tagging. In Georg Rehm, Andreas Witt, and Lothar Lemnitzer, editors, *Data Structures for Linguistic Resources and Applications*. Gunter Narr Verlag, Tübingen, 2007.
- [156] Christian Wartena. Automatic Classification of Social Tags. In Mounia Lalmas, Joemon Jose, Andreas Rauber, Fabrizio Sebastiani, and Ingo Frommholz, editors, *Research and Advanced Technology for Digital Libraries*, volume 6273 of *Lecture Notes in Computer Science*, pages 176–183. Springer Berlin / Heidelberg, 2010.
- [157] Johannes Wheeldon and Jacqueline Faubert. Framing experience: Concept maps, mind maps, and data collection in qualitative research. *International Journal of Qualitative Methods*, 8(3):8–83, 2009.
- [158] Werner Wiater. *Wissensmanagement. Eine Einführung für Pädagogen*. VS Verlag für Sozialwissenschaften, 2007.
- [159] David A. Wiley. *The Instructional Use of Learning Objects*, chapter Connecting learning objects to instructional design theory: a definition, a metaphor, and a taxonomy. The Agency for Instructional Technology, Bloomington, IN, 2002.
- [160] Helmut Wilke. *Systemisches Wissensmanagement*. Lucius und Lucius, 2 edition, 2001.
- [161] Scott Wilson, Oleg Liber, Mark Johnson, Phil Beauvoir, Paul Sharples, and Colin Milligan. Personal Learning Environments: Challenging the dominant design of educational systems. *Journal of e-Learning and Knowledge Society*, 2:27–38, 2007.
- [162] Philip H. Winne, John C. Nesbit, Vive Kumar, Allyson F. Hadwin, Susanne P. Lajoie, Roger Azevedo, and Nancy E. Perry. Supporting self-regulated learning with gStudy software: The Learning Kit Project. *Technology, Instruction, Cognition and Learning*, 3(1/2):105–113, 2006.
- [163] World Wide Web Consortium. SKOS Simple Knowledge Organization System Reference - W3C Recommendation 18 August 2009, 2009. URL <http://www.w3.org/TR/2009/REC-skos-reference-20090818/>. Zuletzt abgerufen am 25.02.2011.
- [164] Jianping Zhang, Da Zhong, and Jiahua Zhang. Knowledge Visualization: An Effective Way of Improving Learning. In *2010 Second International Workshop on Education Technology and Computer Science*, pages 598–601. IEEE, 2010.
- [165] Maayan Zhitomirsky-Geffet, Judit Bar-Ilan, Yitzchak Miller, and Snunith Shoham. A generic framework for collaborative multi-perspective ontology acquisition. *Online Information Review*, 34(1):145–159, 2010.
- [166] Barry J. Zimmerman. A social cognitive view of self-regulated academic learning. *Journal of educational psychology*, 81(3):329–339, 1989.
- [167] Birgit Zimmermann, Christoph Rensing, and Ralf Steinmetz. Making Expert Knowledge of Adaptations of E-Learning Material Available with Patterns. In Klaus Tochtermann and Hermann Maurer, editors, *Proceedings of the 7th International Conference on Knowledge Management*, pages 79–86, Sep 2007.



Abbildungsverzeichnis

2.1	Stufenweise Unterscheidung des Bedeutungsgehaltes	10
2.2	Stufenweise Unterscheidung von Ressource, Information und Wissen	11
2.3	Modell der Wissensspirale	12
2.4	Modell der Bausteine des Wissensmanagements	12
2.5	Münchener Modell	13
2.6	Modell zum persönlichen Wissensmanagement	14
2.7	Wissensreifungsprozess	15
2.8	Bausteine des persönlichen Wissensmanagement	16
2.9	Modell der Prozessschritte für Ressourcen-basiertes Lernen	22
3.1	Akteure und Adressaten beim Tagging	30
3.2	Methoden der Wissensrepräsentation und deren Akteure	31
3.3	Motive für die Nutzung von Tagging-Diensten	33
3.4	Kognitive Analyse des Tagging	34
3.5	Begriffliche Einteilung Wissensrepräsentation, Wissensorganisation und Wissensordnung .	35
3.6	Relationstypen und Wissensorganisationssysteme	36
3.7	Beispiel für ein semantisches Netz mit Typisierung	39
4.1	Beispiel zur Kodierung von semantischen Tags in SKOS	45
4.2	Common Tag-Modell	45
4.3	Tagging Dialog in ZigTag	46
4.4	Tagging Dialog in Faviki	47
4.5	Screenshot von Fuzzzy	48
4.6	Facettierte Navigation im Tag Explorer	49
4.7	Tagging Dialog in Facette	50
4.8	Facettierte Navigation in Facette	51
4.9	Anzahl der typisierten Tags in Evaluationen	56
4.10	Strukturelle Recommendation von Ressourcen	61
5.1	Screenshot von WordNet zum Suchbegriff "Java"	67
5.2	Mapping von Tag-Typen und WordNet-Kategorien im TagExplorer	72
5.3	Vorgehen im ClassTag-System	73
5.4	Vorgehen im ClassTag-System: Tag → WordNet-Kategorie	74
5.5	Verteilung der Tag-Typen in Flickr	74
5.6	Verfahren zur automatischen Erkennung der Tag-Typen	76
6.1	a) Mind-Map, b) Concept-Map	89
6.2	"Miniquire"-Sidebar und "ThingEditor"-Browser	92
6.3	Konzeptuelles Modell einer Personal Learning Environment	94

6.4	Bookmark zur Webseite von Flickr in Delicious	95
6.5	Tag Cloud in Delicious	96
6.6	Browsen in Delicious mittels verwandten Tags	96
6.7	Taggen einer Webseite in Delicious	97
7.1	Basisschema	112
7.2	Erweiterungen im Basisschema	113
7.3	Mehrfaches Speichern einer Webseite	113
7.4	Global genutztes Wissensnetz	114
8.1	Modelliertes Wissensnetz in K-Infinity	118
8.2	Gesamtarchitektur von ELWMS.KOM	120
8.3	Beispiel von Datenaustausch zwischen Client und Server	121
8.4	Sidebar "Recherche"	123
8.5	Alter Tagging-Dialog	124
8.6	Tagging-Dialog	125
8.7	Erweiterungen definieren	126
8.8	Erweiterung im Tagging-Dialog	126
8.9	Zielmanagement	127
8.10	Export der Ziele und Ressourcen als HTML-Seite	128
8.11	Sidebar "Überblick"	129
8.12	Knowledge-Portal	130
A.1	Verfahren zur automatischen Erkennung der Tag-Typen mit Berücksichtigung der URL . .	179
A.2	Verfahren zur automatischen Erkennung der Tag-Typen mit Berücksichtigung von DBpedia	181
A.3	Verfahren zur automatischen Erkennung der Tag-Typen mit Berücksichtigung von DBpedia und URL	183

Tabellenverzeichnis

2.1	Werkzeuge/Hilfsmittel zum Speichern von Rechercheergebnissen	20
4.1	Verwendete Tag-Typen in verwandten Arbeiten	53
4.2	Bedarf der Tag-Typen bei Probanden	57
4.3	Ergebnisse der Evaluation der Tag Recommendation	59
5.1	Beurteilung der Interrater-Reliabilität	69
5.2	Ergebnisse der Klassifizierung von Delicious-Tags vgl. [16]	71
5.3	Anzahl der ausgewählten Tags je Klasse	71
5.4	Ergebnisse der prozentual richtig klassifizierten Tags	72
5.5	Konfusionsmatrix für die Klassifikation der Tag-Typen im Delicious-Korpus	80
5.6	Konfusionsmatrix für die Klassifikation der Tag-Typen im Delicious-Korpus (2/3-Mehrheit)	81
5.7	Konfusionsmatrix für die Klassifikation der Tag-Typen in ELWMS.KOM	83
5.8	Konfusionsmatrix für die Klassifikation der Tag-Typen in ELWMS.KOM (2/3-Mehrheit) . .	83
6.1	Vergleich der verwandten Ansätze hinsichtlich der Anforderungen des Ressourcenmana- gements im Ressourcen-basierten Lernen	101
9.1	Anzahl erstellter Ziele	137
9.2	Anzahl gespeicherter Ressourcen	137
9.3	Allgemein verwendete Hilfsmittel während der Recherche	138
9.4	Unterschiede zwischen Experimental- und Kontrollgruppe in Labor 2	142
9.5	Unterschiede des Zielmanagements zwischen Experimental- und Kontrollgruppe	142
9.6	Unterschiede im Seminar vor Nutzung von ELWMS.KOM und am Ende des Seminars . . .	143
9.7	Meinung der Teilnehmer über ELWMS.KOM nach Ende des Seminars	144
9.8	Usability (SUS Score)	145
9.9	Einschätzung überflüssiger Eingaben in ELWMS.KOM	146
9.10	Wichtigkeit der Funktionen in ELWMS.KOM	147
9.11	Meinung der Teilnehmer über ELWMS.KOM nach Ende der Nutzung	148
9.12	Fragen bzgl. struktureller Desorientierung	148
9.13	Fragen bzgl. konzeptueller Desorientierung	149
9.14	Fragen bzgl. kognitiver Mehrbelastung	149
9.15	Fragen zum Aufwand des Ressourcenmanagements	150
A.1	Konfusionsmatrix für die Klassifikation der Tag-Typen im Delicious-Korpus (nur Eigenna- men und englischsprachige Tags)	175
A.2	Konfusionsmatrix für die Klassifikation der Tag-Typen im Delicious-Korpus (2/3-Mehrheit, nur Eigennamen und englischsprachige Tags)	175
A.3	Konfusionsmatrix für die Klassifikation der Tag-Typen in ELWMS.KOM (nur Eigennamen und englischsprachige Tags)	176

A.4	Konfusionsmatrix für die Klassifikation der Tag-Typen in ELWMS.KOM (2/3-Mehrheit, nur Eigennamen und englischsprachige Tags)	176
A.5	Konfusionsmatrix für die Klassifikation der Tag-Typen im Delicious-Korpus (2/3-Mehrheit und dritte Antwort)	177
A.6	Konfusionsmatrix für die Klassifikation der Tag-Typen in ELWMS (2/3-Mehrheit und dritte Antwort)	177
A.7	Konfusionsmatrix für die Klassifikation der thematischen Ziele in ELWMS.KOM	178
A.8	Konfusionsmatrix für die Klassifikation des Tag-Typs Ziel in ELWMS.KOM (2/3-Mehrheit) .	178
A.9	Konfusionsmatrix für die Klassifikation der Tag-Typen im Delicious-Korpus (mit URL) . . .	179
A.10	Konfusionsmatrix für die Klassifikation der Tag-Typen im Delicious-Korpus (2/3-Mehrheit, mit URL)	180
A.11	Konfusionsmatrix für die Klassifikation der Tag-Typen in ELWMS.KOM (mit URL)	180
A.12	Konfusionsmatrix für die Klassifikation der Tag-Typen in ELWMS.KOM (2/3-Mehrheit, mit URL)	180
A.13	Konfusionsmatrix für die Klassifikation der Tag-Typen im Delicious-Korpus (mit DBpedia)	181
A.14	Konfusionsmatrix für die Klassifikation der Tag-Typen im Delicious-Korpus (2/3-Mehrheit, mit DBpedia)	182
A.15	Konfusionsmatrix für die Klassifikation der Tag-Typen in ELWMS.KOM (mit DBpedia) . . .	182
A.16	Konfusionsmatrix für die Klassifikation der Tag-Typen in ELWMS.KOM (2/3-Mehrheit, mit DBpedia)	182
A.17	Konfusionsmatrix für die Klassifikation der Tag-Typen im Delicious-Korpus (mit DBpedia, mit URL)	183
A.18	Konfusionsmatrix für die Klassifikation der Tag-Typen im Delicious-Korpus (2/3-Mehrheit, mit DBpedia, mit URL)	184
A.19	Konfusionsmatrix für die Klassifikation der Tag-Typen in ELWMS.KOM (mit DBpedia, mit URL)	184
A.20	Konfusionsmatrix für die Klassifikation der Tag-Typen in ELWMS.KOM (2/3-Mehrheit, mit DBpedia, mit URL)	185
B.1	Labor 2 - Rechercheverhalten im Fragebogen vor der Recherche (alle)	187
B.2	Labor 2 - Rechercheverhalten im Fragebogen vor der Recherche (Experimentalgruppe) . .	188
B.3	Labor 2 - Rechercheverhalten im Fragebogen vor der Recherche (Kontrollgruppe)	189
B.4	Labor 2 - Rechercheverhalten im Fragebogen nach der Recherche (alle)	190
B.5	Labor 2 - Rechercheverhalten im Fragebogen nach der Recherche (Experimentalgruppe) .	191
B.6	Labor 2 - Rechercheverhalten im Fragebogen nach der Recherche (Kontrollgruppe)	192
B.7	Einschätzung der Seminarteilnehmer zu ihrem Rechercheverhalten nach der Nutzung von ELWMS.KOM	193
B.8	Einschätzung der Seminarteilnehmer zur Unterstützung von ELWMS.KOM beim Ressourcen-basierten Lernen	195
B.9	Allgemeines Rechercheverhalten in der KOM-Studie	206
B.10	Einschätzung der Teilnehmer der Studie "KOM" zur Unterstützung von ELWMS.KOM beim Ressourcen-basierten Lernen	216

Verzeichnis aller Abkürzungen

API	Application Programming Interface
ELWMS.KOM	E-Learning-Wissens-Management-System
FAQ	Frequently Asked Questions
HTML	Hypertext Markup Language
IEEE	Institute of Electrical and Electronics Engineers
LOM	Learning Object Metadata
LTSC	Learning Technology Standards Consortium
PIMO	Persönliches Informationsmodell
PLE	Personal Learning Environment
RBL	Ressourcen-basiertes Lernen
RDF	Resource Description Framework
RSS	Really Simple Syndication
SCORM	Shareable Content Object Reference Model
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
W3C	World Wide Web Consortium
XESA	Extended Explicit Semantic Analysis



Anhang



A Details zum Verfahren der automatischen Erkennung der Tag-Typen

A.1 Korpus aus Eigennamen und englischsprachigen Tags

a	b	c	d	e	f	← klassifiziert als
88	2	0	0	3	6	a = Person
2	58	0	0	0	0	b = Ort
0	0	2	0	0	0	c = Ereignis
1	0	1	5	2	1	d = Typ
2	1	0	0	63	1	e = Thema
3	0	0	0	8	6	f = Unkategorisiert
0,92	0,95	0,67	1,00	0,83	0,43	Precision
0,89	0,97	1,00	0,50	0,94	0,35	Recall
0,90	0,96	0,80	0,67	0,88	0,39	F-Measure
86,6%						F-Measure Ø

Tabelle A.1: Konfusionsmatrix für die Klassifikation der Tag-Typen im Delicious-Korpus (nur Eigennamen und englischsprachige Tags)

a	b	c	d	e	f	← klassifiziert als
104	2	0	0	4	13	a = Person
3	62	0	0	0	0	b = Ort
0	0	4	0	1	0	c = Ereignis
2	0	2	8	6	1	d = Typ
22	9	1	5	126	11	e = Thema
8	1	5	1	24	16	f = Unkategorisiert
0,75	0,84	0,33	0,57	0,78	0,39	Precision
0,85	0,95	0,80	0,42	0,72	0,29	Recall
0,79	0,89	0,47	0,48	0,75	0,33	F-Measure
71,8%						F-Measure Ø

Tabelle A.2: Konfusionsmatrix für die Klassifikation der Tag-Typen im Delicious-Korpus (2/3-Mehrheit, nur Eigennamen und englischsprachige Tags)

a	b	c	d	e	f	← klassifiziert als
140	1	0	0	2	25	a = Person
0	12	0	0	2	0	b = Ort
0	0	53	0	3	0	c = Ereignis
0	0	0	4	8	0	d = Typ
18	2	1	0	142	45	e = Thema
0	0	0	0	0	0	f = Unkategorisiert
0,89	0,80	0,98	1,00	0,90	0,00	Precision
0,83	0,86	0,95	0,33	0,68	0,00	Recall
0,86	0,83	0,96	0,50	0,78	0,00	F-Measure
82,5%						F-Measure Ø

Tabelle A.3: Konfusionsmatrix für die Klassifikation der Tag-Typen in ELWMS.KOM (nur Eigennamen und englischsprachige Tags)

a	b	c	d	e	f	← klassifiziert als
170	3	1	0	12	31	a = Person
1	13	0	0	2	0	b = Ort
6	0	62	0	4	1	c = Ereignis
0	0	0	6	45	0	d = Typ
34	5	9	4	344	109	e = Thema
0	0	0	0	2	5	f = Unkategorisiert
0,81	0,62	0,86	0,60	0,84	0,03	Precision
0,78	0,81	0,85	0,12	0,68	0,71	Recall
0,79	0,70	0,86	0,20	0,75	0,07	F-Measure
73,3%						F-Measure Ø

Tabelle A.4: Konfusionsmatrix für die Klassifikation der Tag-Typen in ELWMS.KOM (2/3-Mehrheit, nur Eigennamen und englischsprachige Tags)

A.2 Berücksichtigung der Antwort der Nicht-Mehrheit

a	b	c	d	e	f	← klassifiziert als
118	4	0	0	8	14	a = Person
2	146	0	0	8	2	b = Ort
0	0	29	0	0	2	c = Ereignis
2	0	2	27	17	9	d = Typ
40	25	10	3	421	38	e = Thema
17	3	0	1	18	52	f = Unkategorisiert
0,66	0,82	0,71	0,87	0,89	0,44	Precision
0,82	0,92	0,94	0,47	0,78	0,57	Recall
0,73	0,87	0,81	0,61	0,83	0,50	F-Measure
78,2%						F-Measure Ø

Tabelle A.5: Konfusionsmatrix für die Klassifikation der Tag-Typen im Delicious-Korpus (2/3-Mehrheit und dritte Antwort)

a	b	c	d	e	f	← klassifiziert als
172	3	1	0	12	30	a = Person
1	15	0	0	3	0	b = Ort
6	0	62	0	4	1	c = Ereignis
0	5	0	9	41	0	d = Typ
34	7	11	1	394	100	e = Thema
0	0	0	0	0	36	f = Unkategorisiert
0,81	0,50	0,84	0,90	0,87	0,22	Precision
0,79	0,79	0,85	0,16	0,72	1,00	Recall
0,80	0,61	0,84	0,28	0,79	0,35	F-Measure
74,5%						F-Measure Ø

Tabelle A.6: Konfusionsmatrix für die Klassifikation der Tag-Typen in ELWMS (2/3-Mehrheit und dritte Antwort)

A.3 Erkennung von thematischen Zielen

a	b	c	d	e	f	← klassifiziert als
5	0	0	0	0	1	a = Person
0	0	0	0	0	0	b = Ort
0	0	0	0	0	0	c = Ereignis
0	0	0	0	27	0	d = Typ
0	0	9	0	137	28	e = Thema
0	0	0	0	0	0	f = Unkategorisiert
1,00	0,0	0,0	0,0	0,84	0,0	Precision
0,83	0,0	0,0	0,0	0,79	0,0	Recall
0,91	0,0	0,0	0,0	0,81	0,0	F-Measure
70,8%						F-Measure Ø

Tabelle A.7: Konfusionsmatrix für die Klassifikation der thematischen Ziele in ELWMS.KOM

a	b	c	d	e	f	← klassifiziert als
5	0	0	0	0	1	a = Person
0	0	0	0	0	0	b = Ort
0	0	0	0	0	0	c = Ereignis
0	0	0	0	27	0	d = Typ
0	0	9	0	137	28	e = Thema
9	0	0	0	38	18	f = Unkategorisiert
0,36	0,0	0,0	0,0	0,68	0,38	Precision
0,83	0,0	0,0	0,0	0,79	0,28	Recall
0,50	0,0	0,0	0,0	0,73	0,32	F-Measure
55,4%						F-Measure Ø

Tabelle A.8: Konfusionsmatrix für die Klassifikation des Tag-Typs Ziel in ELWMS.KOM (2/3-Mehrheit)

A.4 Berücksichtigung der URL bei der Erkennung des Tag-Typs

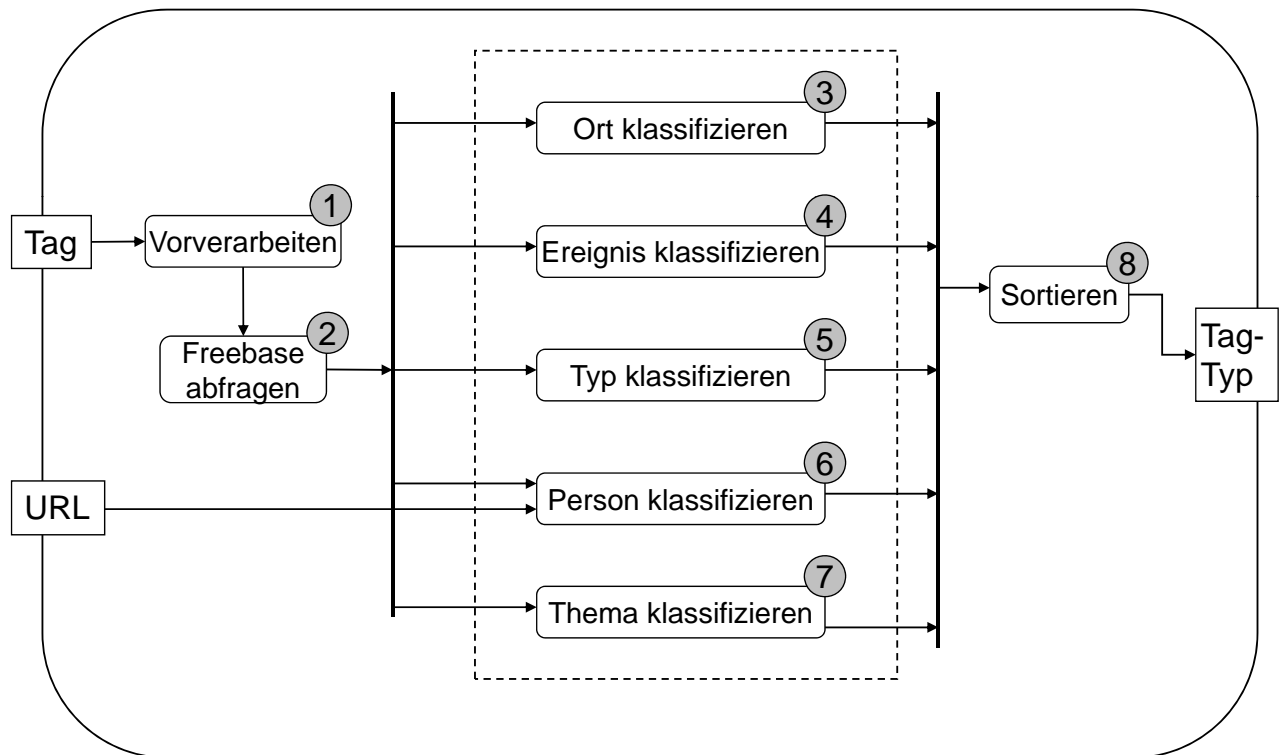


Abbildung A.1: Verfahren zur automatischen Erkennung der Tag-Typen mit Berücksichtigung der URL

a	b	c	d	e	f	← klassifiziert als
99	2	0	0	4	1	a = Person
2	132	0	0	8	2	b = Ort
0	0	15	0	0	1	c = Ereignis
1	0	1	8	9	8	d = Typ
18	16	10	1	228	15	e = Thema
5	0	0	0	9	25	f = Unkategorisiert
0,79	0,88	0,58	0,89	0,88	0,48	Precision
0,93	0,92	0,94	0,30	0,79	0,64	Recall
0,86	0,90	0,71	0,44	0,84	0,55	F-Measure
81,5%						F-Measure Ø

Tabelle A.9: Konfusionsmatrix für die Klassifikation der Tag-Typen im Delicious-Korpus (mit URL)

a	b	c	d	e	f	← klassifiziert als
123	4	0	0	11	3	a = Person
3	146	0	0	8	2	b = Ort
0	0	20	0	3	2	c = Ereignis
8	0	2	19	20	9	d = Typ
53	25	14	11	379	37	e = Thema
25	3	5	1	32	50	f = Unkategorisiert
0,58	0,82	0,49	0,61	0,84	0,49	Precision
0,87	0,92	0,8	0,33	0,73	0,43	Recall
0,70	0,87	0,61	0,43	0,78	0,46	F-Measure
72,1%						F-Measure Ø

Tabelle A.10: Konfusionsmatrix für die Klassifikation der Tag-Typen im Delicious-Korpus (2/3-Mehrheit, mit URL)

a	b	c	d	e	f	← klassifiziert als
144	1	0	0	2	21	a = Person
0	14	0	0	2	0	b = Ort
0	0	53	0	3	0	c = Ereignis
0	5	0	4	10	0	d = Typ
19	2	1	0	160	52	e = Thema
0	0	0	0	0	0	f = Unkategorisiert
0,88	0,64	0,98	1,00	0,90	0,04	Precision
0,86	0,88	0,95	0,21	0,68	0,00	Recall
0,87	0,74	0,96	0,35	0,78	0,00	F-Measure
81,3%						F-Measure Ø

Tabelle A.11: Konfusionsmatrix für die Klassifikation der Tag-Typen in ELWMS.KOM (mit URL)

a	b	c	d	e	f	← klassifiziert als
174	3	1	0	12	27	a = Person
1	15	0	0	3	0	b = Ort
6	0	62	0	4	1	c = Ereignis
0	5	0	6	53	0	d = Typ
44	7	11	4	379	122	e = Thema
1	0	0	0	2	5	f = Unkategorisiert
0,77	0,50	0,84	0,60	0,84	0,03	Precision
0,80	0,79	0,85	0,09	0,67	0,63	Recall
0,79	0,61	0,84	0,16	0,74	0,06	F-Measure
71,3%						F-Measure Ø

Tabelle A.12: Konfusionsmatrix für die Klassifikation der Tag-Typen in ELWMS.KOM (2/3-Mehrheit, mit URL)

A.5 DBpedia als weitere Datenquelle im Verfahren zur Erkennung des Tag-Typs

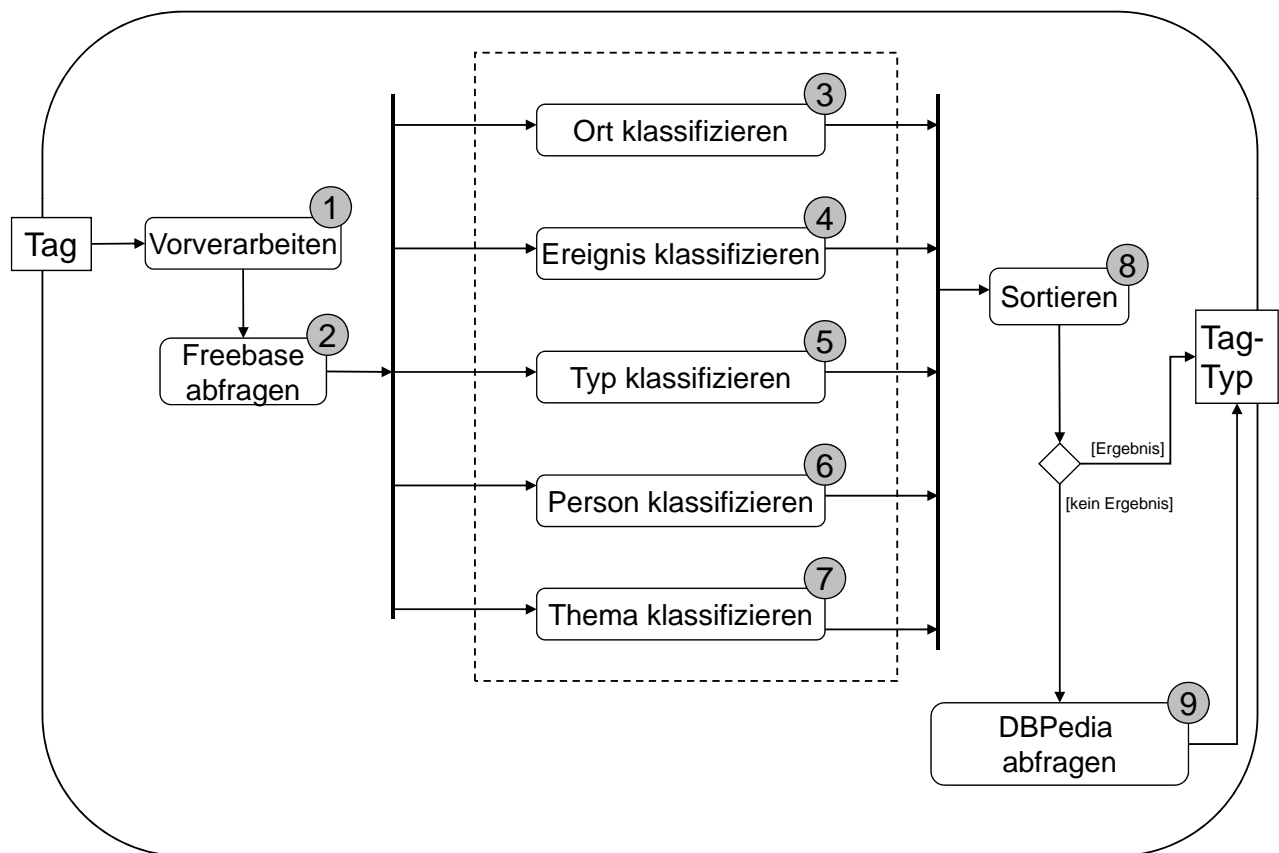


Abbildung A.2: Verfahren zur automatischen Erkennung der Tag-Typen mit Berücksichtigung von DBpedia

a	b	c	d	e	f	← klassifiziert als
91	2	0	0	6	7	a = Person
2	136	0	0	4	2	b = Ort
0	0	15	0	0	1	c = Ereignis
1	0	1	8	9	8	d = Typ
15	17	10	1	230	15	e = Thema
4	0	0	0	11	24	f = Unkategorisiert
0,81	0,88	0,58	0,89	0,88	0,42	Precision
0,86	0,94	0,94	0,30	0,80	0,62	Recall
0,83	0,91	0,71	0,44	0,84	0,50	F-Measure
81,3%						F-Measure Ø

Tabelle A.13: Konfusionsmatrix für die Klassifikation der Tag-Typen im Delicious-Korpus (mit DBpedia)

a	b	c	d	e	f	← klassifiziert als
108	4	0	0	17	12	a = Person
3	150	0	0	4	2	b = Ort
0	0	20	0	3	2	c = Ereignis
2	0	2	19	26	9	d = Typ
43	26	14	11	389	36	e = Thema
22	4	5	1	40	44	f = Unkategorisiert
0,61	0,82	0,49	0,61	0,81	0,42	Precision
0,77	0,94	0,80	0,33	0,75	0,38	Recall
0,68	0,87	0,61	0,43	0,78	0,40	F-Measure
71,2%						F-Measure Ø

Tabelle A.14: Konfusionsmatrix für die Klassifikation der Tag-Typen im Delicious-Korpus (2/3-Mehrheit, mit DBpedia)

a	b	c	d	e	f	← klassifiziert als
140	2	0	0	2	24	a = Person
0	15	0	0	1	0	b = Ort
0	0	53	0	2	1	c = Ereignis
0	5	0	4	10	0	d = Typ
15	2	1	0	178	38	e = Thema
0	0	0	0	0	0	f = Unkategorisiert
0,90	0,63	0,98	1,00	0,92	0,00	Precision
0,83	0,94	0,95	0,21	0,76	0,00	Recall
0,87	0,75	0,96	0,35	0,83	0,00	F-Measure
83,7%						F-Measure Ø

Tabelle A.15: Konfusionsmatrix für die Klassifikation der Tag-Typen in ELWMS.KOM (mit DBpedia)

a	b	c	d	e	f	← klassifiziert als
169	5	1	0	13	29	a = Person
1	16	0	0	2	0	b = Ort
6	0	62	0	4	1	c = Ereignis
0	5	0	6	53	0	d = Typ
33	7	11	3	415	97	e = Thema
1	0	0	0	3	5	f = Unkategorisiert
0,81	0,48	0,84	0,60	0,85	0,04	Precision
0,78	0,84	0,85	0,09	0,73	0,63	Recall
0,79	0,62	0,84	0,16	0,79	0,07	F-Measure
74,0%						F-Measure Ø

Tabelle A.16: Konfusionsmatrix für die Klassifikation der Tag-Typen in ELWMS.KOM (2/3-Mehrheit, mit DBpedia)

A.6 DBpedia als weitere Datenquelle und Berücksichtigung der URL im Verfahren zur Erkennung des Tag-Typs

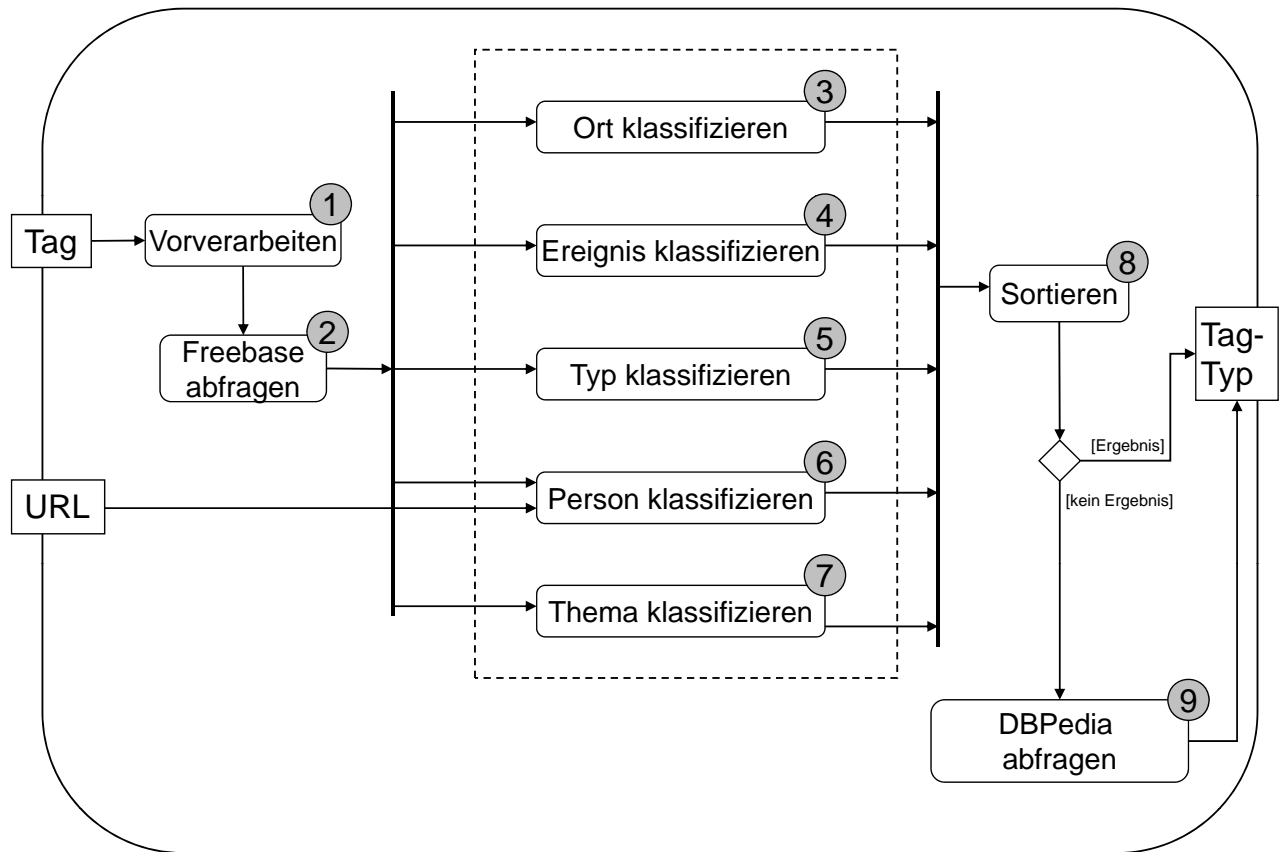


Abbildung A.3: Verfahren zur automatischen Erkennung der Tag-Typen mit Berücksichtigung von DBpedia und URL

a	b	c	d	e	f	← klassifiziert als
99	2	0	0	4	1	a = Person
2	136	0	0	4	2	b = Ort
0	0	15	0	0	1	c = Ereignis
1	0	1	8	9	8	d = Typ
18	17	10	1	227	15	e = Thema
5	0	0	0	10	24	f = Unkategorisiert
0,79	0,88	0,58	0,89	0,89	0,47	Precision
0,93	0,94	0,94	0,30	0,79	0,62	Recall
0,86	0,91	0,71	0,44	0,84	0,53	F-Measure
81,8%						F-Measure Ø

Tabelle A.17: Konfusionsmatrix für die Klassifikation der Tag-Typen im Delicious-Korpus (mit DBpedia, mit URL)

a	b	c	d	e	f	← klassifiziert als
123	4	0	0	11	3	a = Person
3	150	0	0	4	2	b = Ort
0	0	20	0	3	2	c = Ereignis
8	0	2	19	20	9	d = Typ
52	26	14	11	381	35	e = Thema
25	4	5	1	38	43	f = Unkategorisiert
0,58	0,82	0,49	0,61	0,83	0,46	Precision
0,87	0,94	0,80	0,33	0,73	0,37	Recall
0,70	0,87	0,61	0,43	0,78	0,41	F-Measure
71,7%						F-Measure Ø

Tabelle A.18: Konfusionsmatrix für die Klassifikation der Tag-Typen im Delicious-Korpus (2/3-Mehrheit, mit DBpedia, mit URL)

a	b	c	d	e	f	← klassifiziert als
144	1	0	0	2	21	a = Person
0	15	0	0	1	0	b = Ort
0	0	53	0	3	0	c = Ereignis
0	5	0	4	10	0	d = Typ
16	2	1	0	178	37	e = Thema
0	0	0	0	0	0	f = Unkategorisiert
0,90	0,65	0,98	1,00	0,92	0,00	Precision
0,86	0,94	0,95	0,21	0,76	0,00	Recall
0,88	0,77	0,96	0,35	0,83	0,00	F-Measure
84,2%						F-Measure Ø

Tabelle A.19: Konfusionsmatrix für die Klassifikation der Tag-Typen in ELWMS.KOM (mit DBpedia, mit URL)

a	b	c	d	e	f	← klassifiziert als
173	4	1	0	13	26	a = Person
1	16	0	0	2	0	b = Ort
6	0	62	0	4	1	c = Ereignis
0	5	0	6	53	0	d = Typ
41	7	11	4	411	93	e = Thema
1	0	0	0	3	4	f = Unkategorisiert
0,78	0,50	0,84	0,60	0,85	0,30	Precision
0,80	0,84	0,85	0,09	0,72	0,50	Recall
0,79	0,63	0,84	0,16	0,78	0,06	F-Measure
73,6%						F-Measure Ø

Tabelle A.20: Konfusionsmatrix für die Klassifikation der Tag-Typen in ELWMS.KOM (2/3-Mehrheit, mit DBpedia, mit URL)

A.7 Laufzeitverhalten des Verfahrens

Mit einem durchschnittlichen PC (RAM 2GB und CPU 3GHz DualCore) dauert die Bestimmung des Tag-Typs im Durchschnitt 416ms bei einer Standardabweichung von 529ms. Diese Dauer lässt sich durch Optimierung der Implementierung wahrscheinlich noch weiter reduzieren.

Die Daten der externen Quellen wurden für das Verfahren nicht lokal kopiert, sondern direkt bei den Web-Services der externen Quellen abgefragt. Da die Anzahl der möglichen Anfragen bei den externen Quellen häufig begrenzt ist, wurden die Daten lokal zwischengespeichert, um eine mehrfache Wiederholung des Experiments zur Messung der Klassifikationszeit durchzuführen. Der Download der externen Quellen ist für alle verwendeten Quellen möglich, wurde aber auf Grund des Installationsaufwandes vermieden. Die Zeit einer lokalen Datenbankabfrage ist allerdings eher kürzer als die Zeit zum Laden der zwischengespeicherten Daten, die teilweise noch verarbeitet werden müssen.

B Details zu durchgeführten Studien

B.1 [Studie Labor 2] - Auswertung des Fragebogens vor der Recherche

	Immer (6)	Sehr häufig (5)	Häufig (4)	Manch- mal (3)	Selten (2)	Nie (1)	Mittelwert	Standard- abweichung
Ich habe Mühe, Literatur für meine Vorträge und Arbeiten zu finden.	-	1	1	5	1	-	3,25	0,89
Ich plane vor der Recherche, wie ich vorgehen möchte.	1	1	2	3	1	-	3,75	1,28
Ich kann die Qualität und Vertrauenswürdigkeit der gefundenen Webseiten einschätzen.	1	3	2	1	1	-	4,25	1,28
Ich nutze die gefundenen Informationen direkt in meinem Endprodukt (z.B. Hausarbeit, Vortrag) ohne diese extra zu speichern (z.B. in einer Literaturverwaltungs-Software).	2	1	2	-	2	1	3,75	1,91
Ich habe Probleme beim Wiederfinden von Seiten, die ich zu einem früheren Zeitpunkt besucht habe.	-	-	2	3	1	2	2,63	1,19
Ich weiß nicht, wie ich weitersuchen kann, um bessere Suchergebnisse zu bekommen.	-	1	2	4	1	-	3,38	0,92
Ich finde bei der Suche im Internet, wonach ich suche.	-	5	2	1	-	-	4,50	0,76
Mir fällt es schwer zu beurteilen, ob eine Webseite relevant für meine Recherche ist.	-	-	-	4	2	2	2,25	0,89
Ich kann die gefundenen Informationen thematisch einordnen.	3	3	2	-	-	-	5,13	0,83
Mich überfordert die große Menge an Informationen im Internet.	-	1	1	3	1	2	2,75	1,39
Ich führe mit anderen Studenten gemeinsam Recherchen durch.	-	1	-	4	2	1	2,75	1,16
Der Austausch von Recherche-Ergebnissen mit anderen ist umständlich.	1	2	2	1	2	-	3,88	1,46

Tabelle B.1: Labor 2 - Rechercheverhalten im Fragebogen vor der Recherche (alle)

	Immer (6)	Sehr häufig (5)	Häufig (4)	Manch- mal (3)	Selten (2)	Nie (1)	Mittelwert	Standard- abweichung
Ich habe Mühe, Literatur für meine Vorträge und Arbeiten zu finden.	-	-	-	3	1	-	2,75	0,50
Ich plane vor der Recherche, wie ich vorgehen möchte.	1	-	-	2	1	-	3,50	1,73
Ich kann die Qualität und Vertrauenswürdigkeit der gefundenen Webseiten einschätzen.	1	1	1	1	-	-	4,50	0,13
Ich nutze die gefundenen Informationen direkt in meinem Endprodukt (z.B. Hausarbeit, Vortrag) ohne diese extra zu speichern (z.B. in einer Literaturverwaltungs-Software).	2	-	1	-	-	1	4,25	2,36
Ich habe Probleme beim Wiederfinden von Seiten, die ich zu einem früheren Zeitpunkt besucht habe.	-	-	1	-	1	2	2,00	1,41
Ich weiß nicht, wie ich weitersuchen kann, um bessere Suchergebnisse zu bekommen.	-	1	1	2	-	-	3,75	0,96
Ich finde bei der Suche im Internet, wonach ich suche.	-	2	1	1	-	-	4,25	0,96
Mir fällt es schwer zu beurteilen, ob eine Webseite relevant für meine Recherche ist.	-	-	-	2	1	1	2,25	0,96
Ich kann die gefundenen Informationen thematisch einordnen.	1	2	1	-	-	-	5,00	0,82
Mich überfordert die große Menge an Informationen im Internet.	-	-	-	3	-	1	2,50	1,00
Ich führe mit anderen Studenten gemeinsam Recherchen durch.	-	-	-	2	1	1	2,25	0,96
Der Austausch von Recherche-Ergebnissen mit anderen ist umständlich.	1	1	1	-	1	-	4,25	1,71

Tabelle B.2: Labor 2 - Rechercheverhalten im Fragebogen vor der Recherche (Experimentalgruppe)

	Immer (6)	Sehr häufig (5)	Häufig (4)	Manch- mal (3)	Selten (2)	Nie (1)	Mittelwert	Standard- abweichung
Ich habe Mühe, Literatur für meine Vorträge und Arbeiten zu finden.	-	1	1	2	-	-	3,75	0,96
Ich plane vor der Recherche, wie ich vorgehen möchte.	-	1	2	1	-	-	4,00	0,82
Ich kann die Qualität und Vertrauenswürdigkeit der gefundenen Webseiten einschätzen.	-	2	1	-	1	-	4,00	1,41
Ich nutze die gefundenen Informationen direkt in meinem Endprodukt (z.B. Hausarbeit, Vortrag) ohne diese extra zu speichern (z.B. in einer Literaturverwaltungs-Software).	-	1	1	-	2	-	3,25	1,50
Ich habe Probleme beim Wiederfinden von Seiten, die ich zu einem früheren Zeitpunkt besucht habe.	-	-	1	3	-	-	3,25	0,50
Ich weiß nicht, wie ich weitersuchen kann, um bessere Suchergebnisse zu bekommen.	-	-	1	2	1	-	3,00	0,82
Ich finde bei der Suche im Internet, wonach ich suche.	-	3	1	-	-	-	4,75	0,50
Mir fällt es schwer zu beurteilen, ob eine Webseite relevant für meine Recherche ist.	-	-	-	2	1	1	2,25	0,96
Ich kann die gefundenen Informationen thematisch einordnen.	2	1	1	-	-	-	5,25	0,96
Mich überfordert die große Menge an Informationen im Internet.	-	1	1	-	1	1	3,00	1,83
Ich führe mit anderen Studenten gemeinsam Recherchen durch.	-	1	-	2	1	-	3,25	1,26
Der Austausch von Recherche-Ergebnissen mit anderen ist umständlich.	-	1	1	1	1	-	3,50	1,29

Tabelle B.3: Labor 2 - Rechercheverhalten im Fragebogen vor der Recherche (Kontrollgruppe)

B.2 [Studie Labor 2] - Auswertung des Fragebogens nach der Recherche

	Stimme voll und ganz zu (6)	Stimme über- wie- gend zu (5)	Stimme eher zu (4)	Stimme eher nicht zu (3)	Stimme über- wie- gend nicht zu (2)	Stimme über- haupt nicht zu (1)	Mittelwert	Standard- abweichung
Während der Recherche habe ich mir regelmäßig bewusst gemacht, wie ich voran komme und was ich noch machen muss.	3	1	1	3	-	-	5,00	0,93
Während der Recherche habe ich mir Mühe gegeben.	2	2	4	-	-	-	4,88	0,99
Während der Bearbeitung der Aufgabe war ich genervt.	-	-	4	-	3	1	2,88	1,25
Ich habe mich angestrengt, weil ich in der Aufgabe gut abschneiden wollte.	2	4	1	-	1	-	4,75	1,28
Ich habe konzentriert gearbeitet.	2	3	1	1	1	-	4,50	1,41
Ich fand es schwierig, Zusammenhänge zwischen den Informationen auf den unterschiedlichen Webseiten herzustellen.	1	-	1	4	1	1	3,13	1,46
Während der Aufgabe habe ich mich gestresst gefühlt.	1	2	2	-	-	3	3,38	2,07
Mit meiner Arbeitsweise bin ich zufrieden.	1	2	3	1	1	-	4,13	1,25
Während der Recherche habe ich mir Sorgen gemacht, ob ich die Aufgabe schaffe.	2	1	2	-	2	1	3,75	1,91
Während der Recherche habe ich mich wohl gefühlt.	-	2	5	1	-	-	4,13	0,64
Ich bin mit dem Ergebnis dieser Aufgabe zufrieden.	-	1	3	3	1		3,50	0,93
Während der Recherche hatte ich Probleme, das Gelesene mit dem zu verbinden, was ich schon wusste.	-	1	1	2	-	4	2,38	1,60
Die Bearbeitung der Aufgabe hat mir Spaß gemacht.	1	1	3	3	-	-	4,00	1,07
Die Rechercheaufgabe hat mich überfordert.	-	-	2	2	2	2	2,50	1,20
Die Bearbeitungszeit für diese Aufgabe war zu kurz.	1	1	-	3	2	1	3,13	1,64

Tabelle B.4: Labor 2 - Rechercheverhalten im Fragebogen nach der Recherche (alle)

	Stimme voll und ganz zu (6)	Stimme über- wie- gend zu (5)	Stimme eher zu (4)	Stimme eher nicht zu (3)	Stimme über- wie- gend nicht zu (2)	Stimme über- haupt nicht zu (1)	Mittelwert	Standard- abweichung
Während der Recherche habe ich mir regelmäßig bewusst gemacht, wie ich voran komme und was ich noch machen muss.	3	1	-	-	-	-	5,75	0,50
Während der Recherche habe ich mir Mühe gegeben.	2	1	1	-	-	-	5,25	0,96
Während der Bearbeitung der Aufgabe war ich nervt.	-	-	2	-	2	-	3,00	1,15
Ich habe mich angestrengt, weil ich in der Aufgabe gut abschneiden wollte.	2	2	-	-	-	-	5,50	0,58
Ich habe konzentriert gearbeitet.	2	2	-	-	-	-	5,50	0,58
Ich fand es schwierig, Zusammenhänge zwischen den Informationen auf den unterschiedlichen Webseiten herzustellen.	1	-	-	1	1	1	3,00	2,16
Während der Aufgabe habe ich mich gestresst gefühlt.	1	-	2	-	-	1	3,75	2,06
Mit meiner Arbeitsweise bin ich zufrieden.	-	2	1	1	-	-	4,25	0,96
Während der Recherche habe ich mir Sorgen gemacht, ob ich die Aufgabe schaffe.	1	-	2	-	-	1	3,75	2,06
Während der Recherche habe ich mich wohl gefühlt.	-	1	3	-	-	-	4,25	0,50
Ich bin mit dem Ergebnis dieser Aufgabe zufrieden.	-	1	2	-	1	-	3,75	1,26
Während der Recherche hatte ich Probleme, das Gelesene mit dem zu verbinden, was ich schon wusste.	-	-	1	1	-	2	2,25	1,50
Die Bearbeitung der Aufgabe hat mir Spaß gemacht.	1	1	1	1	-	-	4,50	1,29
Die Rechercheaufgabe hat mich überfordert.	-	-	1	1	1	1	2,50	1,29
Die Bearbeitungszeit für diese Aufgabe war zu kurz.	1	-	-	2	-	1	3,25	2,06

Tabelle B.5: Labor 2 - Rechercheverhalten im Fragebogen nach der Recherche (Experimentalgruppe)

	Stimme voll und ganz zu (6)	Stimme über- wie- gend zu (5)	Stimme eher zu (4)	Stimme eher nicht zu (3)	Stimme über- wie- gend nicht zu (2)	Stimme über- haupt nicht zu (1)	Mittelwert	Standard- abweichung
Während der Recherche habe ich mir regelmäßig bewusst gemacht, wie ich voran komme und was ich noch machen muss.	-	-	1	3	-	-	4,25	0,50
Während der Recherche habe ich mir Mühe gegeben.	-	1	3	-	-	-	4,50	1,00
Während der Bearbeitung der Aufgabe war ich genervt.	-	-	2	-	1	1	2,75	1,50
Ich habe mich angestrengt, weil ich in der Aufgabe gut abschneiden wollte.	-	2	1	-	1	-	4,00	1,41
Ich habe konzentriert gearbeitet.	-	1	1	1	1	-	3,50	1,29
Ich fand es schwierig, Zusammenhänge zwischen den Informationen auf den unterschiedlichen Webseiten herzustellen.	-	-	1	3	-	-	3,25	0,50
Während der Aufgabe habe ich mich gestresst gefühlt.	-	2	-	-	-	2	3,00	2,31
Mit meiner Arbeitsweise bin ich zufrieden.	1	-	2	-	1	-	4,00	1,63
Während der Recherche habe ich mir Sorgen gemacht, ob ich die Aufgabe schaffe.	1	1	-	-	2	-	3,75	2,06
Während der Recherche habe ich mich wohl gefühlt.	-	1	2	1	-	-	4,00	0,82
Ich bin mit dem Ergebnis dieser Aufgabe zufrieden.	-	-	1	3	-	-	3,25	0,50
Während der Recherche hatte ich Probleme, das Gelesene mit dem zu verbinden, was ich schon wusste.	-	1	-	1	-	2	2,50	1,91
Die Bearbeitung der Aufgabe hat mir Spaß gemacht.	-	-	2	2	-	-	3,50	0,58
Die Rechercheaufgabe hat mich überfordert.	-	-	1	1	1	1	2,50	1,29
Die Bearbeitungszeit für diese Aufgabe war zu kurz.	-	1	-	1	2	-	3,00	1,41

Tabelle B.6: Labor 2 - Rechercheverhalten im Fragebogen nach der Recherche (Kontrollgruppe)

B.3 [Studie Seminar 2] - Auswertung des Fragebogens nach der Nutzung von ELWMS.KOM

	Stimme voll und ganz zu (6)	Stimme über- wie- gend zu (5)	Stimme eher zu (4)	Stimme eher nicht zu (3)	Stimme über- wie- gend nicht zu (2)	Stimme über- haupt nicht zu (1)	Mittelwert	Standard- abweichung
Vor den Recherchen habe ich mir überlegt, wie ich bei der Recherche vorgehen möchte.	1	4	2	-	-	-	4,86	0,69
Ich habe Mühe, Literatur für meine Vorträge und Arbeiten zu finden.	1	3	2	-	1	-	4,43	1,27
Ich kann die Qualität und Vertrauenswürdigkeit der gefundenen Webseiten einschätzen.	1	1	4	1	-	-	4,29	0,95
Ich habe die gefundenen Informationen direkt in meiner Seminararbeit genutzt	1	2	2	1	1	-	4,14	1,35
Ich finde bei der Suche im Internet, wonach ich suche.	2	2	3	-	-	-	4,86	0,90
Während der Recherchen habe ich mir regelmäßig bewusst gemacht, wie ich voran komme und was ich noch machen muss.	3	2	1	1	-	-	5,00	1,15
Ich habe Probleme beim Wiederfinden von Seiten, die ich zu einem früheren Zeitpunkt besucht habe.	1	-	1	-	4	1	2,71	1,70
Während der Recherchen habe ich mir Mühe gegeben.	3	2	1	1	-	-	5,00	1,15
Ich weiß nicht, wie ich weitersuchen kann, um bessere Suchergebnisse zu bekommen.	-	2	1	3	1	-	3,57	1,13
Während der Bearbeitung der Seminararbeit war ich wegen ELWMS genervt.	-	-	1	3	2	1	2,57	0,98
Ich habe mich angestrengt, weil ich in der Seminararbeit gut abschneiden wollte.	5	1	1	-	-	-	5,57	0,79
Während der Recherchen habe ich gemerkt, dass meine Gedanken abgeschweift sind.	-	1	1	3	2	-	3,14	1,07
Während der Recherchen hatte ich Probleme, das Gelesene mit dem zu verbinden, was ich schon weiß.	-	-	1	2	3	1	2,43	0,98
Mir fällt es schwer zu beurteilen, ob eine Webseite relevant für meine Recherche ist.	-	-	-	2	4	1	2,14	0,69
Ich kann die gefundenen Informationen thematisch einordnen.	-	6	1	-	-	-	4,86	0,38
Ich habe konzentriert gearbeitet.	3	-	3	1	-	-	4,71	1,25
Ich fand es schwierig, Zusammenhänge zwischen den Informationen auf den unterschiedlichen Webseiten herzustellen.	-	-	1	3	2	1	2,57	0,98
Mich hat die große Menge an Informationen überfordert, die ich beim Recherchieren im Internet finden kann.	-	1	1	1	4	-	2,86	1,21
Während der Seminararbeit habe ich mich bei den Recherchen gestresst gefühlt	-	-	3	2	2	-	3,14	0,90
Mit meiner Arbeitsweise bin ich zufrieden.	-	3	4	-	-	-	4,43	0,53
Während der Recherchen habe ich mir Sorgen gemacht, ob ich die Aufgabe schaffe.	-	2	1	1	3	-	3,29	1,38
Während der Recherchen habe ich mich wohl gefühlt.	-	1	3	3	-	-	3,71	0,76
Ich bin mit dem Ergebnis der Recherchen für die Seminararbeit zufrieden.	-	5	2	-	-	-	4,71	0,49

Tabelle B.7: Einschätzung der Seminarteilnehmer zu ihrem Rechercheverhalten nach der Nutzung von ELWMS.KOM

	Stimme voll und ganz zu (6)	Stimme über- wie- gend zu (5)	Stimme eher zu (4)	Stimme eher nicht zu (3)	Stimme über- wie- gend nicht zu (2)	Stimme über- haupt nicht zu (1)	Mittelwert	Standard- abweichung
Die Möglichkeit Ziele anzulegen hat mich bei der Planung meiner Recherchen unterstützt.	-	2	2	1	1	1	3,43	1,51
Wenn eine Webseite bereits von anderen Nutzern gespeichert wurde, hat mich dies bei der Einschätzung der Qualität und Vertrauenswürdigkeit der Webseite unterstützt.	-	1	2	1	2	1	3,00	1,41
Durch die Möglichkeit, gefundene Informationen direkt mit der Sidebar abzuspeichern, konnte ich Wechsel zwischen Browser und anderen Anwendungen reduzieren bzw. vermeiden.	3	1	2	1	-	-	4,86	1,21
Das Taggen mit typisierten Tags hat mir geholfen, Seiten wiederzufinden, die ich zu einem früheren Zeitpunkt bereits besucht habe.	1	2	3	-	-	1	4,14	1,57
Die Vorschläge in der Sidebar und im Portal haben mir geholfen, weitere relevante Informationen zu finden.	2	-	2	3	-	-	4,14	1,35
Ich habe im Wissensnetz gefunden, wonach ich gesucht habe.	1	1	4	1	-	-	4,29	0,95
Die Planung der Recherche hat mir während der Durchführung bei der Beurteilung der Relevanz der Webseiten geholfen.	-	-	3	2	1	-	3,33	0,82
Beim Taggen habe ich mir überlegt zu welchem Thema ich die Webseite einsortieren kann.	4	1	1	1	-	-	5,14	1,21
Während einer Recherche habe ich mich auf das Lesen der gefundenen Informationen konzentriert.	-	2	4	-	-	1	3,86	1,35
Ich konnte von den Webseiten und Wissensnetzen der anderen Nutzer profitieren.	-	1	2	2	1	1	3,14	1,35
Der Austausch von Recherche-Ergebnissen mit anderen Nutzern von ELWMS war umständlich.	-	-	-	3	2	-	2,60	0,55
Die eingeblendete Sidebar hat die Gefahr reduziert, dass ich mich von anderen Webseiten ablenken lasse, die für die aktuelle Recherche nicht relevant sind.	-	-	3	1	2	1	2,86	1,21
Wenn ich an das Wissensnetz denke, das während meiner Recherche entstanden ist, finde ich den Aufwand des Abspeicherns der Webseiten angemessen.	-	-	4	3	-	-	3,57	0,53
Tags, die beim Tagging vorgeschlagen wurden, haben das Tagging gut unterstützt.	2	1	3	1	-	-	4,57	1,13
Ich finde das Wissensnetz, das bei mir entstanden ist, unübersichtlich.	-	-	4	3	-	-	3,57	0,53
Beim Browsen im Wissensnetz habe ich neue Zusammenhänge entdeckt.	-	-	2	3	1	1	2,86	1,07
Das Zurechtfinden in Wissensnetzen anderer Nutzer war schwierig.	-	1	4	-	2	-	3,57	1,13
Die Ziele in der Sidebar haben mir geholfen, eine begonnene Recherche weiter zu planen und wenn nötig meine Suchstrategie anzupassen.	-	-	3	1	2	1	2,86	1,21

	Stimme voll und ganz zu (6)	Stimme über- wie- gend zu (5)	Stimme eher zu (4)	Stimme eher nicht zu (3)	Stimme über- wie- gend nicht zu (2)	Stimme über- haupt nicht zu (1)	Mittelwert	Standard- abweichung
Die Ziele und deren zugeordnete Webseiten haben mich dabei unterstützt, mich an eine frühere oder unterbrochene Recherche zu erinnern und schnelle Anknüpfungspunkte zu haben.	-	-	2	4	-	1	3,00	1,00
Die typisierten Tags gaben mir die Möglichkeit, Zusammenhänge zwischen einer gefunden Webseite und der Recherche gut im Wissensnetz abzuspeichern.	1	2	3	-	-	1	4,14	1,57
Das Umordnen im Wissensnetz und Einfügen neuer Informationen ist einfach.	-	1	3	2	1	-	3,57	0,98

Tabelle B.8: Einschätzung der Seminarteilnehmer zur Unterstützung von ELWMS.KOM beim Ressourcen-basierten Lernen

B.4 [Studie KOM] - Fragebogen vor der Nutzung von ELWMS.KOM

Herzlich Willkommen bei der Evaluation von ELWMS.

ELWMS steht kurz für E-Learning-Wissens-Management-System. Dieses System wurde im Rahmen meiner Dissertation entwickelt. Ich freue mich, dass Sie an der Evaluation dieses Systems teilnehmen. Bitte füllen Sie diesen Fragebogen aus bevor Sie ELWMS das erste Mal nutzen. Am Ende der Befragung werden Sie auf die Webseite mit Tutorial und Download umgeleitet. Lesen Sie sich die Fragen bitte genau durch und beantworten Sie diese spontan. Ich werde Informationen zu Ihrer Person, zur Benutzung des Computers bei Recherchen, Ihrem Vorgehen bei Recherchen und Ihre Erfahrungen mit Web 2.0 erheben.

Demographische Angaben

1. Identifikationsnummer

2. Geschlecht

☐ weiblich ☐ männlich

3. Alter:

4. Muttersprache

☐ Deutsch
☐ Sonstiges _____

5. Studieren Sie?

☐ Ja (weiter mit Frage 6)
☐ Nein (überspringen Sie die Fragen 7-10)

6. Bitte füllen Sie die Textfelder anhand ihrer aktuellen Tätigkeit aus.

Ich arbeite als _____
bei _____

7. Bitte geben Sie den von Ihnen angestrebten Abschluss an.

☐ Bachelor
☐ Master
☐ Diplom
☐ Magister
☐ Sonstiges

8. Semesterzahl im derzeitigen Studiengang

9. Waren Sie vor Ihrem derzeitigen Studiengang bereits in einem anderen Studiengang eingeschrieben?

- ☐ Ja, mit Abschluss _____
- ☐ Ja, ohne Abschluss _____
- ☐ Nein

10. Haben Sie vor dem Studium eine Berufsausbildung gemacht?

- ☐ Ja, mit Abschluss _____
- ☐ Ja, ohne Abschluss _____
- ☐ Nein

Nutzung des Computers für Recherchezwecke

11. Wie viele Stunden benutzen Sie den Computer wöchentlich im Durchschnitt (Job/Studium und Freizeit)?

12. Wie oft surfen Sie im Internet?

- ☐ mehrmals am Tag mehrmals am Tag
- ☐ einmal am Tag einmal am Tag
- ☐ 3-5 Tage in der Woche 3-5 Tage in der Woche
- ☐ 1-2 Tage in der Woche 1-2 Tage in der Woche
- ☐ alle paar Wochen oder weniger alle paar Wochen oder weniger

13. Wie häufig recherchieren Sie für Ihr Studium / Ihren Job (im Internet, in der Bibliothek,...)?

- ☐ oft
- ☐ manchmal
- ☐ selten
- ☐ nie

14. Wonach suchen Sie, wenn Sie für Ihr Studium / Ihren Job recherchieren?

- ☐ Hintergrundinformationen für einen Vortrag / (Seminar-)Arbeit / Publikation
- ☐ Aktuelle Informationen aus Politik, Wirtschaft & Gesellschaft
- ☐ Fakten (Jahreszahlen, Namen, etc.)
- ☐ Wissenschaftliche Publikationen
- ☐ Zeitschriftenartikel

- ☐ Lehrmittel und Unterrichtsmaterialien
- ☐ Lehrbücher (z.B. für eigene Prüfungsvorbereitung)
- ☐ Illustrationsmaterial für Vorträge (Bilder, Videos) / (Seminar-)Arbeit / Publikation
- ☐ Software
- ☐ Sonstiges: _____

15. Wie häufig benutzen Sie das Internet für Ihre Recherchen?

- ☐ immer
- ☐ sehr häufig
- ☐ häufig
- ☐ manchmal
- ☐ selten
- ☐ nie

16. Wann haben Sie das erste Mal eine Recherche im Internet durchgeführt?

- ☐ Schule
- ☐ Berufsausbildung
- ☐ Freizeit
- ☐ Studium
- ☐ Job

17. Vor wie vielen Jahren war diese erste Recherche?

18. Sind Sie mit dem Webbrowser Mozilla Firefox vertraut?

- ☐ ja
- ☐ nein
- ☐ unsicher

19. Welchen Webbrowser verwenden Sie üblicherweise?

- ☐ Internet Explorer
- ☐ Opera
- ☐ Google Chrome
- ☐ Mozilla Firefox
- ☐ Andere: _____

20. Wo suchen Sie, wenn Sie eine Recherche im Internet durchführen? Schätzen Sie in % wie sich die Anteile der einzelnen Suchmöglichkeiten auf die Recherchen aufteilen. (Beispielsweise 50% Google, 30% Google Scholar, 20% Wikipedia.) Summe aller Einträge muss gleich 100 sein.

Suchmaschine (z.B. Google, Yahoo, Bing,..)	_____
Google Scholar	_____
Wikipedia	_____
Bibliothekskatalog	_____
Portal	_____
Fachdatenbank	_____
Sonstiges	_____

Bisheriges Rechercheverhalten

21. Bitte geben Sie an, inwieweit Sie den folgenden Aussagen zustimmen.

	Stimme voll und ganz zu	Stimme über- wie- gend zu	Stimme eher zu	Stimme eher nicht zu	Stimme über- wie- gend nicht zu	Stimme über- haupt nicht zu
Ich habe Mühe, Literatur für meine Vorträge und Arbeiten zu finden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich plane vor einer Recherche, wie ich vorgehen möchte.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich kann die Qualität und Vertrauenswürdigkeit der gefundenen Webseiten einschätzen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich nutze die gefundenen Informationen direkt in meinem Endprodukt (z.B. Paper, Hausarbeit, Vortrag) ohne diese extra zu speichern (z.B. in einer Literaturverwaltungs-Software).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich habe Probleme beim Wiederfinden von Seiten, die ich zu einem früheren Zeitpunkt besucht habe.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich weiß nicht, wie ich weitersuchen kann, um bessere Suchergebnisse zu bekommen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich finde bei der Suche im Internet, wonach ich suche.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	Stimme voll und ganz zu	Stimme über- wie- gend zu	Stimme eher zu	Stimme eher nicht zu	Stimme über- wie- gend nicht zu	Stimme über- haupt nicht zu
Mir fällt es schwer zu beurteilen, ob eine Webseite relevant für meine Recherche ist.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich kann die gefundenen Informationen thematisch einordnen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mich überfordert die große Menge an Informationen, die ich beim Recherchieren im Internet finden kann.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich führe mit anderen Personen (Kollegen/Studenten) gemeinsam Recherchen durch.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Der Austausch von Recherche-Ergebnissen mit anderen ist umständlich.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich lasse mich während einer Recherche ablenken und lese Seiten, die für die eigentliche Recherche nicht relevant sind.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn ich recherchiere und gefundene Informationen lese, tue ich viele Dinge nebenher.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bin damit zufrieden, welchen Aufwand das Abspeichern von Webseiten verursacht, wenn ich mir das Ergebnis nach einigen Recherchen anschau.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich würde gern von den Informationen profitieren, die meine Kollegen oder andere Personen zusammentragen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Während der Recherchen habe ich Probleme, das Gelesene mit dem zu verbinden, was ich schon weiß.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich finde es schwierig, Zusammenhänge zwischen den Informationen auf den unterschiedlichen Webseiten herzustellen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich passe meine Suchstrategie während einer Recherche an.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	Stimme voll und ganz zu	Stimme über- wie- gend zu	Stimme eher zu	Stimme eher nicht zu	Stimme über- wie- gend nicht zu	Stimme über- haupt nicht zu
Eine Recherche fortzusetzen nach ich sie unterbre- chen musste, fällt mir schwer. Ich weiß dann nicht, welche Webseiten ich schon besucht habe und wie ich weiter machen wollte.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Beim Speichern einer Webseite halte ich auch den Kontext der Recherche fest.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich ändere die Struktur nicht (z.B. Ordner auf der Festplatte), in der ich Webseiten speichere, weil dies umständlich ist.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich habe Spaß daran, neue Software auszuprobie- ren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Für mich ist es nützlich, eine Software zur Unter- stützung meiner Recherchen einzusetzen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bin motiviert, ELWMS für meine nächsten Re- cherchen einzusetzen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**22. Beschreiben Sie kurz, wie Sie bei einer Recherche im Internet üblicherweise vorgehen, bei-
spielsweise bei der letzten Recherche.**

23. Welche Funktionen nutzen / kennen Sie in Ihrem Browser?

	Nutze ich	Kenne ich, nutze ich aber nicht	Kenne ich nicht
Lesezeichen / Bookmarks	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Chronik / History / Verlauf	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tab-Funktion (Öffnen von Webseiten in Tabs in einem Fenster)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sitzungsmanager (Beim nächsten Öffnen des Browsers werden die zuletzt geöffneten Webseiten wieder geladen)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Erweiterungen (Add-ons) (Erweiterungen sind z.B. Zotero, Newsreader-Plugin, Web Developer)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

24. Falls Sie Erweiterungen wie z.B. Zotero, Newsreader etc. nutzen, welche sind das?

25. Gibt es weitere Funktionen im Browser, die Sie nutzen? Wenn ja, welche sind das? Wenn nein, schreiben Sie bitte "Keine".

26. Welche Software / Hilfsmittel nutzen Sie, zusätzlich zum Browser, regelmäßig für Ihre Recherchen im Internet, um z.B. Ihre Suchergebnisse zu speichern?

	Nutze ich	Kenne ich, nutze ich aber nicht	Kenne ich nicht
Ordner auf der Festplatte	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
E-Mail	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Dokumente (z.B. Word, NotePad)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tabellen (z.B. in Excel, Word)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Notizen auf Papier	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Literaturverwaltungs-Software (z.B. Jabref)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Online-Bookmarking Dienste (z.B. Delicious)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mind-Mapping-Tools	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

27. Falls Sie eine Literaturverwaltungs-Software bei der Recherche nutzen, welche Software nutzen Sie?

28. Falls Sie einen Online-Bookmarking Dienst bei der Recherche nutzen, welchen Dienst nutzen Sie?

29. Gibt es weitere Hilfsmittel, die Sie für Ihre Recherche nutzen? Wenn ja, welche sind das? Wenn nein, schreiben Sie bitte "Keine".

Tagging

30. Wissen Sie, was man unter dem Begriff Tagging versteht?

- ☐ Ja
☐ Nein (weiter mit Frage 36)

31. Nutzen Sie bereits die Möglichkeit des Taggings?

- ☐ Ja
☐ Nein (weiter mit Frage 36)

32. Wie lange nutzen Sie Tagging schon?

33. Wofür nutzen Sie das Tagging?

- ☐ Bilder
☐ Videos
☐ Bookmarks
☐ Andere: _____

34. Haben Sie bestimmte Tagging-Konventionen?

- ☐ Ja, ich habe eigene Tagging-Konventionen entwickelt
☐ Nein, ich tagge eher chaotisch

35. Bitte beschreiben Sie Ihre Art zu taggen.

36. Welche der folgenden Aktionen haben Sie schon durchgeführt?

	Ja	Nein	Kenne ich nicht/ ich weiß nicht, was gemeint ist
Fotos in einen Bilderdienst (z.B. Flickr) hochgeladen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Videos in einen Videodienst (z.B. YouTube) hochgeladen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Eigenen Blog/Blogbeitrag erstellt	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Eine Seite in einem Wiki (z.B. Wikipedia) bearbeitet/erstellt	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Einen Online-Bookmarking Dienst (z.B. Delicious) genutzt, um Informationen zu finden, die andere Benutzer dort gesammelt haben	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mit einer Blog-Suchmaschine (z.B. Technorati) gezielt nach Blogs gesucht	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tag Clouds genutzt, um durch Informationen zu navigieren.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

-
37. Vielen Dank für das Ausfüllen dieses Fragebogens. Wenn es noch etwas gibt, was Sie mir mitteilen wollen, können Sie gerne das folgende Feld nutzen. Nach Klick auf Absenden werden Sie auf eine Webseite mit einem Tutorial zu ELWMS und dem Download weitergeleitet und können mit dem Erstellen Ihres Wissensnetzes loslegen. In einigen Wochen werde ich Sie bitten, einen Fragebogen zur Einschätzung von ELWMS auszufüllen.

B.5 [Studie KOM] - Auswertung des Fragebogens vor der Nutzung von ELWMS.KOM

	Stimme voll und ganz zu (6)	Stimme über- wie- gend zu (5)	Stimme eher zu (4)	Stimme eher nicht zu (3)	Stimme über- wie- gend nicht zu (2)	Stimme über- haupt nicht zu (1)	Mittelwert	Standard- abweichung
Ich habe Mühe, Literatur für meine Vorträge und Arbeiten zu finden.	2%	9%	9%	26%	34%	19%	2,62	1,27
Ich plane vor einer Recherche, wie ich vorgehen möchte.	0%	19%	19%	32%	23%	8%	3,19	1,21
Ich kann die Qualität und Vertrauenswürdigkeit der gefundenen Webseiten einschätzen.	13%	51%	21%	8%	8%	0%	4,55	1,07
Ich nutze die gefundenen Informationen direkt in meinem Endprodukt (z.B. Paper, Hausarbeit, Vortrag) ohne diese extra zu speichern (z.B. in einer Literaturverwaltungs-Software).	2%	19%	15%	25%	19%	21%	2,98	1,46
Ich habe Probleme beim Wiederfinden von Seiten, die ich zu einem früheren Zeitpunkt besucht habe.	6%	19%	28%	25%	17%	6%	3,55	1,29
Ich weiß nicht, wie ich weitersuchen kann, um bessere Suchergebnisse zu bekommen.	4%	9%	11%	17%	49%	9%	2,74	1,29
Ich finde bei der Suche im Internet, wonach ich suche.	6%	57%	30%	4%	4%	0%	4,57	0,82
Mir fällt es schwer zu beurteilen, ob eine Webseite relevant für meine Recherche ist.	0%	9%	9%	26%	40%	15%	2,58	1,15
Ich kann die gefundenen Informationen thematisch einordnen.	17%	55%	17%	11%	0%	0%	4,77	0,87
Mich überfordert die große Menge an Informationen, die ich beim Recherchieren im Internet finden kann.	9%	4%	32%	25%	21%	9%	3,28	1,36
Ich führe mit anderen Personen (Kollegen/Studenten) gemeinsam Recherchen durch.	0%	9%	19%	17%	36%	19%	2,64	1,26
Der Austausch von Recherche-Ergebnissen mit anderen ist umständlich.	6%	25%	25%	17%	21%	8%	3,55	1,41
Ich lasse mich während einer Recherche ablenken und lese Seiten, die für die eigentliche Recherche nicht relevant sind.	13%	28%	28%	15%	9%	6%	4,04	1,37
Wenn ich recherchiere und gefundene Informationen lese, tue ich viele Dinge nebenher.	11%	19%	38%	15%	9%	8%	3,85	1,36
Ich bin damit zufrieden, welchen Aufwand das Abspeichern von Webseiten verursacht, wenn ich mir das Ergebnis nach einigen Recherchen anschau.	4%	15%	32%	25%	21%	4%	3,45	1,20
Ich würde gern von den Informationen profitieren, die meine Kollegen oder andere Personen zusammentragen.	45%	32%	11%	8%	4%	0%	5,08	1,11
Während der Recherchen habe ich Probleme, das Gelesene mit dem zu verbinden, was ich schon weiß.	2%	6%	8%	25%	49%	11%	2,53	1,10
Ich finde es schwierig, Zusammenhänge zwischen den Informationen auf den unterschiedlichen Webseiten herzustellen.	0%	15%	11%	23%	40%	11%	2,79	1,25

	Stimme voll und ganz zu (6)	Stimme überwiegend zu (5)	Stimme eher zu (4)	Stimme eher nicht zu (3)	Stimme überwiegend nicht zu (2)	Stimme überhaupt nicht zu (1)	Mittelwert	Standardabweichung
Ich passe meine Suchstrategie während einer Recherche an.	13%	51%	21%	11%	4%	0%	4,58	0,99
Eine Recherche fortzusetzen nach ich sie unterbrechen musste, fällt mir schwer. Ich weiß dann nicht, welche Webseiten ich schon besucht habe und wie ich weiter machen wollte.	8%	11%	30%	26%	19%	6%	3,45	1,29
Beim Speichern einer Webseite halte ich auch den Kontext der Recherche fest.	0%	9%	11%	21%	38%	21%	2,51	1,22
Ich ändere die Struktur nicht (z.B. Ordner auf der Festplatte), in der ich Webseiten speichere, weil dies umständlich ist.	6%	19%	28%	11%	19%	17%	3,30	1,54
Ich habe Spaß daran, neue Software auszuprobieren	8%	42%	32%	15%	2%	2%	4,32	1,01
Für mich ist es nützlich, eine Software zur Unterstützung meiner Recherchen einzusetzen.	19%	32%	26%	11%	9%	2%	4,34	1,29
Ich bin motiviert, ELWMS für meine nächsten Recherchen einzusetzen.	32%	32%	21%	9%	4%	2%	4,74	1,23

Tabelle B.9: Allgemeines Rechercheverhalten in der KOM-Studie

B.6 [Studie KOM] - Fragebogen nach der Nutzung von ELWMS.KOM

Herzlich Willkommen bei der Evaluation von ELWMS.ELWMS steht kurz für E-Learning-Wissens-Management-System. Dieses System wurde im Rahmen meiner Dissertation entwickelt. Ich freue mich, dass Sie ELWMS genutzt haben. Bitte füllen Sie diesen Fragebogen aus nachdem Sie sich mit ELWMS vertraut gemacht haben. Lesen Sie sich die Fragen bitte genau durch und beantworten Sie diese spontan. Ich werde Informationen zu Benutzung des Computers bei Recherchen und Ihre Erfahrungen mit ELWMS erheben.

Allgemeine Nutzung von ELWMS

1. Identifikationsnummer

2. Wie häufig haben Sie ELWMS für Ihre Recherchen genutzt?

- ☐ Immer
☐ Sehr häufig
☐ Ab und zu
☐ Selten
☐ Nie

3. Wie viele Minuten haben Sie im Durchschnitt täglich im Internet recherchiert?

4. Bitte geben Sie an, inwieweit Sie den folgenden Aussagen zustimmen.

	Stimme voll und ganz zu	Stimme eher zu	Stimme teils teils zu	Stimme eher nicht zu	Stimme über- haupt nicht zu
Ich denke, dass ich dieses System gerne häufig nutzen würde.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich fand das System unnötig komplex.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich denke, das System war einfach zu benutzen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich denke, dass ich die Hilfe eines Technikers benötigen würde, um das System benutzen zu können.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich halte die verschiedenen Funktionen des Systems für gut integriert.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	Stimme voll und ganz zu	Stimme eher zu	Stimme teils teils zu	Stimme eher nicht zu	Stimme überhaupt nicht zu
Ich halte das System für zu inkonsistent.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich kann mir vorstellen, dass die meisten Leute sehr schnell lernen würden, mit dem System umzugehen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich fand das System sehr mühsam zu benutzen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich fühlte mich bei der Nutzung des Systems sehr sicher.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich musste viele Dinge lernen, bevor ich das System nutzen konnte.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

5. Bitte geben Sie an, inwieweit Sie den folgenden Aussagen zustimmen.

	Weiß nicht	Stimme voll und ganz zu	Stimme überwiegend zu	Stimme eher zu	Stimme eher nicht zu	Stimme überwiegend nicht zu	Stimme überhaupt nicht zu
Die Möglichkeit Ziele anzulegen hat mich bei der Planung meiner Recherchen unterstützt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Wenn eine Webseite bereits von anderen Nutzern gespeichert wurde, hat mich dies bei der Einschätzung der Qualität und Vertrauenswürdigkeit der Webseite unterstützt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Durch die Möglichkeit, gefundene Informationen direkt mit der Sidebar abzuspeichern, konnte ich Wechsel zwischen Browser und anderen Anwendungen reduzieren bzw. vermeiden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	Weiß nicht	Stimme voll und ganz zu	Stimme über- wie- gend zu	Stimme eher zu	Stimme eher nicht zu	Stimme über- wie- gend nicht zu	Stimme über- haupt nicht zu
Das Taggen mit typisierten Tags hat mir geholfen, Seiten wiederzufinden, die ich zu einem früheren Zeitpunkt bereits besucht habe.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Vorschläge in der Sidebar und im Portal haben mir geholfen, weitere relevante Informationen zu finden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich habe im Wissensnetz gefunden, wonach ich gesucht habe.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Planung der Recherche hat mir während der Durchführung bei der Beurteilung der Relevanz geholfen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Beim Taggen habe ich mir überlegt zu welchem Thema ich die Webseite einsortieren kann.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Während einer Recherche konzentriere ich mich auf das Lesen der gefundenen Informationen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich konnte von den Webseiten und Wissensnetzen der anderen Nutzer profitieren.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Der Austausch von Recherche-Ergebnissen mit anderen Nutzern von ELWMS war umständlich.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die eingeblendete Sidebar hat die Gefahr reduziert, dass ich mich von anderen Webseiten ablenken lasse, die für die aktuelle Recherche nicht relevant sind.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	Weiß nicht	Stimme voll und ganz zu	Stimme über- wie- gend zu	Stimme eher zu	Stimme eher nicht zu	Stimme über- wie- gend nicht zu	Stimme über- haupt nicht zu
Wenn ich an das Wissensnetz denke, das während meiner Recherche entstanden ist, finde ich den Aufwand des Abspeicherns der Webseiten angemessen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tags, die beim Tagging vorgeschlagen wurden, haben das Tagging gut unterstützt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich finde das Wissensnetz, das bei mir entstanden ist, unübersichtlich.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Beim Browsen im Wissensnetz habe ich neue Zusammenhänge entdeckt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Das Zurechtfinden in Wissensnetzen anderer Nutzer war schwierig.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Ziele in der Sidebar haben mir geholfen, eine begonnene Recherche weiter zu planen und wenn nötig meine Suchstrategie anzupassen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Ziele und deren zugeordnete Webseiten haben mich dabei unterstützt, mich an eine frühere oder unterbrochene Recherche zu erinnern und schnelle Anknüpfungspunkte zu haben.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die typisierten Tags gaben mir die Möglichkeit, Zusammenhänge zwischen einer gefundenen Webseite und der Recherche gut im Wissensnetz abzuspeichern.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	Weiß nicht	Stimme voll und ganz zu	Stimme über- wie- gend zu	Stimme eher zu	Stimme eher nicht zu	Stimme über- wie- gend nicht zu	Stimme über- haupt nicht zu
Das Umordnen im Wissensnetz und Einfügen neuer Informationen ist einfach.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

6. Bitte geben Sie an, inwieweit Sie den folgenden Aussagen zustimmen.

	Unbedingt notwen- dig	Ab und zu sinnvoll	Kein Bedarf	Ich weiß nicht, welche Funktion gemeint ist
Speichern von markierten/ausgewählten Abschnitten	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Taggen von Webseiten	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tags, die beim Tagging vorgeschlagen werden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Angabe des Typs des Tags (z.B. Thema, Person,...)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Taggen von Tags (Verknüpfung von Tags untereinander, z.B. Person - Thema)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Aufbau einer Themenstruktur (Unterthema/Oberthema)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Möglichkeit, sich Ziele zu setzen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Zuordnung von Webseiten zu Zielen (Webseiten mit Zielen taggen bzw. Ziel aktivieren [Button mit lila Punkt], um Webseiten automatisch zuzuordnen)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Anzeige in der Sidebar, ob eine Webseite bereits gespeichert ist	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Weitere Webseiten werden in der Sidebar vorgeschlagen (z.B. Webseiten, die zu einem ähnlichen Thema abgespeichert wurden)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	Unbedingt notwen- dig	Ab und zu sinnvoll	Kein Bedarf	Ich weiß nicht, welche Funktion gemeint ist
Ansicht von Webseiten und Tags im Portal (Klick auf den Home-Button/ Auge, um in das Portal zu gelangen)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Graphisch Browsen im Portal (Netzdarstellung)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Suche im Portal nach Webseiten und Tags	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Möglichkeit in den Wissensnetzen (Ressourcen, Tags) anderer Benutzer zu browsen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Erweiterungen zu definieren, um mehr Informationen abspeichern zu können (eigene Formularfelder beim Speichern ähnlich dem BibTex-Feld)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Das Exportieren des Wissensnetzes als BibTex / HTML	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

7. Hatten Sie das Gefühl, bei der Arbeit mit ELWMS Eingaben gemacht zu haben, welche Ihnen eigentlich überflüssig erschienen?

- ☐ Keine überflüssigen Eingaben
- ☐ Wenige überflüssige Eingaben
- ☐ Einige überflüssige Eingaben
- ☐ Viele überflüssige Eingaben

8. Falls Ihnen Eingaben als überflüssig erschienen: Welche waren das?

9. Haben Sie sich im Werkzeug zusätzliche Funktionen zur Bearbeitung der Aufgabe gewünscht?

- ☐ Ja
- ☐ Nein (weiter mit Frage 11)

10. Falls ja: Welche sind das?

11. Kreuzen Sie an, welche Tag-Typen Sie für Ihre Recherchen nützlich finden.

- ☐ Thema
- ☐ Ziel
- ☐ Person
- ☐ Ort
- ☐ Ereignis
- ☐ Typ

12. Gibt es Tag-Typen, die wichtig für Ihre Recherchen sind aber in ELWMS fehlen?

13. Sollen die Tags automatisch typisiert werden (beispielsweise: wenn Sie "Ralf Steinmetz" als Tag eingeben, dass ELWMS zuverlässig automatisch erkennt, dass dies eine Person ist oder "Darmstadt" als Ort erkennt)?

- ☐ Ja
- ☐ Nein

14. Wie würden Sie vorgehen, wenn Sie nach einer Weile eine Webseite im Wissensnetz wiederfinden wollen. Bitte beschreiben Sie.

Zukünftige Nutzung von ELWMS

15. Sie haben ELWMS genutzt, um Recherchen durchzuführen. Bitte geben Sie an, inwieweit Sie den folgenden Aussagen zustimmen.

	Stimme voll und ganz zu	Stimme über- wie- gend zu	Stimme eher zu	Stimme eher nicht zu	Stimme über- wie- gend nicht zu	Stimme über- haupt nicht zu
Ich habe ELWMS für meine Recherchen während meiner Recherchen gern genutzt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ELWMS war sehr gut auf die Anforderungen der Recherche zugeschnitten.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ELWMS hat mir geholfen, bei der Recherche strukturierter vorzugehen als normalerweise.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich würde ELWMS meinen Freunden empfehlen, wenn diese Recherchen durchführen und ihre Webseiten verwalten müssen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
ELWMS war mir bei der Organisation der gefundenen Quellen (Webseiten) nützlich.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Qualität meiner Recherchen schätze ich durch die Nutzung von ELWMS besser ein.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich würde ELWMS gern für weitere Recherchen nutzen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

16. Möchten Sie uns noch etwas mitteilen?

B.7 [Studie KOM] - Auswertung des Fragebogens nach der Nutzung von ELWMS.KOM

	Stimme voll und ganz zu (6)	Stimme überwiegend zu (5)	Stimme eher zu (4)	Stimme eher nicht zu (3)	Stimme überwiegend nicht zu (2)	Stimme überhaupt nicht zu (1)	Weiß nicht	Mittelwert	Standardabweichung
Die Möglichkeit Ziele anzulegen hat mich bei der Planung meiner Recherchen unterstützt.	25%	25%	15%	20%	5%	0%	10%	4,50	1,29
Wenn eine Webseite bereits von anderen Nutzern gespeichert wurde, hat mich dies bei der Einschätzung der Qualität und Vertrauenswürdigkeit der Webseite unterstützt.	5%	25%	30%	5%	0%	0%	35%	4,46	0,78
Durch die Möglichkeit, gefundene Informationen direkt mit der Sidebar abzuspeichern, konnte ich Wechsel zwischen Browser und anderen Anwendungen reduzieren bzw. vermeiden.	60%	15%	5%	10%	10%	0%	0%	5,05	1,43
Das Taggen mit typisierten Tags hat mir geholfen, Seiten wiederzufinden, die ich zu einem früheren Zeitpunkt bereits besucht habe.	30%	25%	15%	15%	10%	0%	5%	4,53	1,39
Die Vorschläge in der Sidebar und im Portal haben mir geholfen, weitere relevante Informationen zu finden.	0%	10%	40%	15%	5%	10%	20%	3,44	1,21
Ich habe im Wissensnetz gefunden, wonach ich gesucht habe.	0%	15%	10%	30%	10%	15%	20%	3,00	1,37
Die Planung der Recherche hat mir während der Durchführung bei der Beurteilung der Relevanz geholfen.	20%	15%	20%	15%	5%	10%	15%	4,00	1,66
Beim Taggen habe ich mir überlegt zu welchem Thema ich die Webseite einsortieren kann.	30%	50%	10%	0%	0%	10%	0%	4,80	1,44
Während einer Recherche konzentriere ich mich auf das Lesen der gefundenen Informationen.	20%	35%	20%	10%	10%	5%	0%	4,30	1,45
Ich konnte von den Webseiten und Wissensnetzen der anderen Nutzer profitieren.	0%	5%	40%	10%	5%	20%	20%	3,06	1,39
Der Austausch von Recherche-Ergebnissen mit anderen Nutzern von ELWMS war umständlich.	0%	10%	10%	20%	15%	5%	40%	3,08	1,24
Die eingeblendete Sidebar hat die Gefahr reduziert, dass ich mich von anderen Webseiten ablenken lasse, die für die aktuelle Recherche nicht relevant sind.	20%	25%	10%	5%	15%	10%	15%	4,00	1,80
Wenn ich an das Wissensnetz denke, das während meiner Recherche entstanden ist, finde ich den Aufwand des Abspeicherns der Webseiten angemessen.	15%	50%	5%	5%	5%	0%	20%	4,81	1,05

	Stimme voll und ganz zu (6)	Stimme über- wie- gend zu (5)	Stimme eher zu (4)	Stimme eher nicht zu (3)	Stimme über- wie- gend nicht zu (2)	Stimme über- haupt nicht zu (1)	Weiß nicht	Mittelwert	Standard- abweichung
Tags, die beim Tagging vorgeschlagen wurden, haben das Tagging gut unterstützt.	15%	20%	30%	5%	5%	0%	25%	4,47	1,13
Ich finde das Wissensnetz, das bei mir entstanden ist, unübersichtlich.	5%	0%	40%	5%	15%	10%	25%	3,27	1,39
Beim Browsen im Wissensnetz habe ich neue Zusammenhänge entdeckt.	0%	20%	20%	20%	5%	0%	35%	3,85	0,99
Das Zurechtfinden in Wissensnetzen anderer Nutzer war schwierig.	5%	15%	5%	15%	10%	0%	50%	3,80	1,40
Die Ziele in der Sidebar haben mir geholfen, eine begonnene Recherche weiter zu planen und wenn nötig meine Suchstrategie anzupassen.	10%	45%	5%	10%	10%	0%	20%	4,44	1,26
Die Ziele und deren zugeordnete Webseiten haben mich dabei unterstützt, mich an eine frühere oder unterbrochene Recherche zu erinnern und schnelle Anknüpfungspunkte zu haben.	15%	45%	5%	10%	10%	0%	15%	4,53	1,28
Die typisierten Tags gaben mir die Möglichkeit, Zusammenhänge zwischen einer gefundenen Webseite und der Recherche gut im Wissensnetz abzuspeichern.	20%	40%	20%	10%	0%	5%	5%	4,58	1,26
Das Umordnen im Wissensnetz und Einfügen neuer Informationen ist einfach.	10%	20%	20%	20%	5%	0%	25%	4,13	1,19

Tabelle B.10: Einschätzung der Teilnehmer der Studie "KOM" zur Unterstützung von ELWMS.KOM beim Ressourcen-basierten Lernen

C Schriftenverzeichnis

C.1 Hauptveröffentlichungen

1. Doreen Böhnstedt, Christoph Rensing, and Ralf Steinmetz. Supporting Resource Based Learning with Knowledge Networks and Scaffolds. In Andreas Kaminski, Max Mühlhäuser, Werner Sesink, and Jürgen Steimle, editors, *Interdisciplinary Approaches to Technology Enhanced Learning. Interdisziplinäre Zugänge zu technologie-gestütztem Lernen. [in Vorbereitung]*. Waxmann, Münster 2011.
2. Doreen Böhnstedt, Philipp Scholl, Christoph Rensing, and Ralf Steinmetz. Enhancing an Environment for Knowledge Acquisition based on Web Resources by Automatic Tag Type Identification. In Michael E. Auer and Jeanne Schreurs, editors, *Proceedings of International Conference on Computer-aided Learning 2010 (ICL 2010)*, Seiten 380–389. Kassel University Press 2010.
3. Doreen Böhnstedt, Philipp Scholl, Christoph Rensing, and Ralf Steinmetz. Collaborative Semantic Tagging of Web Resources on the Basis of Individual Knowledge Networks. In Geert-Jan Houben, Gord McCalla, Fabio Pianesi, and Massimo Zancanaro, editors, *Proceedings of First and Seventeenth International Conference on User Modeling, Adaptation, and Personalization UMAP 2009*, volume 5535 of *Lecture Notes in Computer Science*, Seiten 379–384. Springer-Verlag Berlin Heidelberg 2009.
4. Doreen Böhnstedt, Philipp Scholl, Christoph Rensing, and Ralf Steinmetz. Modeling Personal Knowledge Networks to Support Resource Based Learning. In Klaus Tochtermann and Hermann Maurer, editors, *Proceedings of 9th International Conference on Knowledge Management and Knowledge Technologies (I-KNOW'09)*, Journal of Universal Computer Science, Seiten 309–316. Verlag der Technischen Universität Graz, Austria, Universiti Malaysia Sarawak, Malaysia, and Know-Center, Austria 2009.
5. Doreen Böhnstedt, Philipp Scholl, Bastian Benz, Christoph Rensing, Ralf Steinmetz, and Bernhard Schmitz. Einsatz persönlicher Wissensnetze im Ressourcen-basierten Lernen. In Silke Seehusen, Ulrike Lucke, and Stefan Fischer, editors, *DeLFI 2008: 6. e-Learning Fachtagung Informatik*, P-132, Seiten 113–124. GI, Lecture Notes in Informatics (LNI), Köllen, Bonn 2008.
6. Doreen Böhnstedt, Philipp Scholl, Christoph Rensing, and Ralf Steinmetz. ELWMS.KOM - Typisiertes Tagging in persönlichen Wissensnetzen. In Ulrike Lucke, Martin Christof Kindsmüller, Stefan Fischer, Michael Herczeg, and Silke Seehusen, editors, *Workshop Proceedings der Tagungen Mensch & Computer 2008, DeLFI 2008 und Cognitive Design 2008*, Seiten 330–331. Logos Verlag, Berlin 2008.
7. Doreen Mann, Philipp Scholl, Christoph Rensing, and Ralf Steinmetz. Interaktive, Community-unterstützte Wissensnetze für persönliches Wissensmanagement in Ressourcen-basierten Lernkontexten. In Christoph Rensing and Guido Rößling, editors, *Proceedings der Pre-Conference Workshops der 5. e-Learning Fachtagung Informatik - DeLFI 2007*, Seiten 35–42. Logos Verlag, Berlin 2007.

C.2 Mitautorenschaft und sonstige Veröffentlichungen

8. Philipp Scholl, Doreen Böhnstedt, Christoph Rensing, and Ralf Steinmetz. Adaptivity and Adaptability in Personal and Community Knowledge Networks. In Andreas Kaminski, Max Mühlhäuser, Werner Sesink, and Jürgen Steimle, editors, *Interdisciplinary Approaches to Technology Enhanced Learning. Interdisziplinäre Zugänge zu technologie-gestütztem Lernen. [in Vorbereitung]*. Waxmann, Münster 2011.
9. Christoph Rensing, Philipp Scholl, Doreen Böhnstedt, and Ralf Steinmetz. Recommending and Finding Multimedia Resources in Knowledge Acquisition Based on Web Resources. In *Proceedings of 19th International Conference on Computer Communications and Networks*, Seiten 1–6. IEEE, IEEE eXpress Conference Publishing 2010.
10. Philipp Scholl, Doreen Böhnstedt, Renato Domínguez García, Christoph Rensing, and Ralf Steinmetz. Extended Explicit Semantic Analysis for Calculating Semantic Relatedness of Web Resources. In Martin Wolpers, Paul A. Kirschner, Maren Scheffel, Stefanie Lindstädt, and Vania Dimitrova, editors, *Sustaining TEL: From Innovation to Learning and Practice Proceedings of EC-TEL 2010*, volume 6383 of *Lecture Notes in Computer Science*, Seiten 324–339. Springer Verlag 2010.
11. Renato Domínguez García, Alexandru Berlea, Philipp Scholl, Doreen Böhnstedt, Christoph Rensing, and Ralf Steinmetz. Improving Topic Exploration in the Blogosphere by Detecting Relevant Segments. In Klaus Tochtermann and Hermann Maurer, editors, *Proceedings of 9th International Conference on Knowledge Management and Knowledge Technologies (I-KNOW'09)*, Journal of Universal Computer Science, Seiten 177–188. Verlag der Technischen Universität Graz, Austria, Universiti Malaysia Sarawak, Malaysia, and Know-Center, Austria 2009.
12. Renato Domínguez García, Doreen Böhnstedt, Philipp Scholl, Christoph Rensing, and Ralf Steinmetz. Von Tags zu semantischen Netzen - Einsatz im Ressourcen-basierten Lernen. In Andreas Schwill and Nicolas Apostopoulos, editors, *Lernen im digitalen Zeitalter - Workshop-Band - Dokumentation der Pre-Conference zur DeLFI 2009*, Seiten 29–36. Logos, Berlin 2009.
13. Christoph Rensing and Doreen Böhnstedt. Bildung einer Community zur Vermittlung von E-Learning-Erfahrungen auf Basis semantischer Netze. In Andreas Schwill and Nicolas Apostopoulos, editors, *Lernen im digitalen Zeitalter - Proceedings der 7. e-Learning Fachtagung Informatik der Gesellschaft für Informatik, Berlin 2009*, 153, Seiten 187–198. Gesellschaft für Informatik, GI Verlag, Bonn 2009.
14. Philipp Scholl, Bastian F. Benz, Doreen Böhnstedt, Christoph Rensing, Bernhard Schmitz, and Ralf Steinmetz. Implementation and Evaluation of a Tool for Setting Goals in Self-Regulated Learning with Web Resources. In Ulrike Cress, Vania Dimitrova, and Marcus Specht, editors, *Learning in the Synergy of Multiple Disciplines, EC-TEL 2009*, volume LNCS Vol 5794, Seiten 521–534. Springer-Verlag Berlin Heidelberg 2009.
15. Philipp Scholl, Doreen Böhnstedt, Renato Dominguez Garcia, Christoph Rensing, and Ralf Steinmetz. Anwendungen und Nutzen der Automatischen Erkennung von Web-Genres in persönlichen

und Community- Wissensnetzen. In Andreas Schwill and Nicolas Apostopoulos, editors, *Lernen im digitalen Zeitalter - Workshop-Band - Dokumentation der Pre-Conference zur DeLFI 2009*, Seiten 37–44. Logos, Berlin 2009.

16. Philipp Scholl, Renato Dominguez Garcia, Doreen Böhnstedt, Christoph Rensing, and Ralf Steinmetz. Towards Language-Independent Web Genre Detection. In ACM, editor, *WWW '09: Proceedings of the 18th international conference on World wide web*, Seiten 1157–1158. ACM, Madrid, Spain 2009.
17. Renato Dominguez Garcia, Philipp Scholl, Doreen Böhnstedt, Christoph Rensing, and Ralf Steinmetz. Towards To an Automatic Web Genre Classification. Technical Report TR-2008-10, Multimedia Communications Lab, Darmstadt 2008.
18. Philipp Scholl, Bastian Benz, Doreen Böhnstedt, Christoph Rensing, Ralf Steinmetz, and Bernhard Schmitz. Einsatz und Evaluation eines Zielmanagement-Werkzeugs bei der selbstregulierten Internet-Recherche. In Silke Seehusen, Ulrike Lucke, and Stefan Fischer, editors, *DeLFI 2008: 6. e-Learning Fachtagung Informatik*, P-132, Seiten 125–136. GI, Lecture Notes in Informatics (LNI), Köllen, Bonn 2008.
19. Philipp Scholl, Bastian Benz, Doreen Mann, Christoph Rensing, Ralf Steinmetz, and Bernhard Schmitz. Scaffolding von selbstreguliertem Lernen in einer Rechercheumgebung für internetbasierte Ressourcen. In Christoph Rensing and Guido Rößling, editors, *Proceedings der Pre-Conference Workshops der 5. e-Learning Fachtagung Informatik - DeLFI 2007*, Seiten 43–50. Logos Verlag, Berlin 2007.
20. Philipp Scholl, Doreen Mann, Christoph Rensing, and Ralf Steinmetz. Support of Acquisition and Organization of Knowledge Artifacts in Informal Learning Contexts. In *European Distance and E-Learning Network: EDEN - Book of Abstracts*, Seiten 16–17. Naples 2007.



D Lebenslauf des Verfassers

Persönliche Daten

Name:	Doreen Böhnstedt
Geburtsdatum:	06. Oktober 1981
Geburtsort:	Eilenburg
Nationalität:	Deutsch
Familienstand:	verheiratet

Akademischer Werdegang

seit 09/2006	Technische Universität Darmstadt, Fachgebiet Multimedia Kommunikation – Beginn der Promotion
2006 – 2009	Technische Universität Darmstadt, Graduiertenkolleg "Qualitätsverbesserung im E-Learning durch rückgekoppelte Prozesse" – Stipendium
2004 – 2006	BTU Cottbus – Studium der Informations- und Medientechnik, Abschluss: Master of Science
2000 – 2004	BTU Cottbus – Studium der Informations- und Medientechnik, Abschluss: Bachelor of Science
1992 – 2000	Karl-August-Möbius-Gymnasium in Eilenburg, Abschluss: Abitur

Berufliche Tätigkeiten

seit 09/2009	Technische Universität Darmstadt, Fachgebiet Multimedia Kommunikation – Wissenschaftliche Mitarbeiterin
04/2007 – 08/2009	Technische Universität Darmstadt, Fachgebiet Multimedia Kommunikation – Wissenschaftliche Hilfskraft
05/2005 – 08/2006	Bayliss Graphics – Freie Mitarbeiterin im Bereich Webdesign und -entwicklung
01/2005 – 04/2005	Tequima Europe GmbH – Praktikum
08/2002 – 09/2002	Medienwerkstatt JB – Praktikum
07/2001 – 06/2004	Schülerhilfe – Durchführung von Nachhilfekursen in Mathematik, Deutsch und Englisch

Tätigkeiten in der Lehre

Seit SS 2007	Technische Universität Darmstadt, Assistentin für die Vorlesung und Übung "Einführung in Net Centric Systems"
Seit 09/2006	Technische Universität Darmstadt, Betreuerin für verschiedene Studien-, Bachelor-, Master- und Diplomarbeiten

Darmstadt, den 12. April 2011

E Erklärung laut §9 der Promotionsordnung

Ich versichere hiermit, dass ich die vorliegende Dissertation allein und nur unter Verwendung der angegebenen Literatur verfasst habe. Die Arbeit hat bisher noch nicht zu Prüfungszwecken gedient.

Darmstadt, 2011

Doreen Böhnstedt
